



Descifrando las bases genéticas de enfermedades complejas: Enfermedad de Behçet

Lourdes Ortiz Fernández

Tesis doctoral

Sevilla, 2015

Memoria presentada por Lourdes Ortiz Fernández para optar al grado de Doctor por la Universidad de Sevilla.

El presente trabajo de investigación ha sido realizado en el Instituto de Biomedicina de Sevilla (IBiS), bajo la supervisión de la Doctora María Francisca González Escribano.

Lourdes Ortiz Fernández ha recibido financiación del Programa de Ayudas Predoctorales de Formación en Investigación en Salud, pFIS (11/00547) del Instituto de Salud Carlos III.

El proyecto de investigación ha recibido financiación Fondo de Investigaciones sanitarias (10/1701 y 13/01118) del Instituto de Salud Carlos III, y la Consejería de Salud de la Junta de Andalucía (PI0411/2010).



D.^a M^a FRANCISCA GONZÁLEZ ESCRIBANO, DOCTORA EN
BIOLOGÍA Y FACULTATIVA ESPECIALISTA DE ÁREA EN
INMUNOLOGÍA EN EL HOSPITAL VIRGEN DEL ROCÍO DE SEVILLA

CERTIFICA

que el trabajo de investigación que lleva por título “Descifrando las bases genéticas de enfermedades complejas: Enfermedad de Behçet” ha sido realizado por la licenciada D.^a Lourdes Ortiz Fernández bajo su dirección y que reúne las condiciones necesarias para ser leído y defendido como tesis para optar al grado de Doctor por la Universidad de Sevilla.

En Sevilla, a 2 de Octubre de 2015

Fdo.: M^a Francisca González Escribano

Directora de tesis

Agradecimientos

En primer lugar quiero agradecer a mi directora de tesis, Dra. M^a Francisca González Escribano, por acompañarme y guiarme durante los cuatro años en los que ha transcurrido este proyecto de investigación. Paca, gracias por ser un referente, por tu completa dedicación y por transmitirme un gran entusiasmo por la ciencia. Pero sobre todo, te tengo que agradecer la confianza depositada en mí, traducida en la energía necesaria para llegar a encontrarme hoy aquí, escribiendo los agradecimientos de mi tesis doctoral.

Gracias al Dr. Antonio Núñez Roldán, por brindarme la oportunidad de formar parte de su equipo. Gracias por inspirar en nosotros los valores de un buen docente y la importancia de transmitir el conocimiento. Al resto de facultativos del Servicio de Inmunología, en especial al Dr. Marco Antonio Montes Cano y al Dr. José Raúl García Lozano por enriquecer el trabajo con sus aportaciones científicas y por su total disponibilidad, así como a la Dra. Isabel Aguilera por su apoyo.

A la Dra. Marta Conde, compañera de viaje y pilar fundamental para que este trabajo se haya llevado a cabo. Amiga, gracias por el apoyo y por tener siempre palabras de ánimo y energía positiva. A Elena Aguado, gracias por tu bondad y generosidad, dispuesta siempre a ayudar, siempre optimista y con una sonrisa. Ha sido todo un privilegio ser vuestra compañera y amiga, hacemos un buen equipo.

Gracias a mis primeros compañeros de laboratorio, Mariajo, José Luis, María José, Mari Carmen y Fuen. Gracias por acogerme con los brazos abiertos y hacerme sentir como en casa desde el primer momento que cruce la puerta del laboratorio. Ha sido una gran suerte encontrar tan buenos amigos en compañeros de laboratorio.

A todos los compañeros del Servicio de Inmunología, en especial a José Manuel, Ángel y Cristina por esos buenos ratos que hemos pasado juntos, tanto dentro como fuera del laboratorio. A las compañeras de la sección de tipaje y de biología molecular por su paciencia y por enseñarme siempre con alegría.

A los compañeros y técnicos de los servicios del IBiS por sus consejos y ayuda prestada durante estos años. Gracias al resto de compañeros y amigos, gracias a esa multidisciplinariedad de grupo que hace que siempre encuentres la solución a cualquier

problema. A Álvaro, gracias por tus consejos y ayuda con las imágenes y el diseño, y gracias junto con Pablo por el equipo de meriendas, que ayudan a soportar los días que se alargan. A Bea y Jessi, grandes investigadoras y compañeras de piso, gracias por el apoyo, sobre todo durante estos últimos meses. A Ismael, por su paciencia, tranquilidad y por recordarme que hay que perseguir los sueños. A Estefanía, Marina, Victor, Enriqueta, Rafa, Luis, Ana... gracias por tan buenos momentos que hemos pasado juntos.

Al Dr. Javier Martín por abrirme las puertas del laboratorio cuando todavía era estudiante y por ser partícipe de mis primeros pasos en el mundo de la investigación. Desde entonces, siempre ha estado cerca de mi camino investigador, acoguéndome en su laboratorio en múltiples ocasiones. Gracias a todo su equipo, en especial a David por enseñarme con dedicación y por su paciencia en las largas conversaciones telefónicas resolviéndome dudas, pero sobre todo, gracias por la confianza que siempre ha depositado en mí. Agradecer también a Lara por estar siempre disponible a ayudar.

A mis biólogos granadinos, que aunque la mayoría hayan emigrado de Granada, me hacen llegar su cariño desde múltiples países e idiomas y gracias por ofrecerme tantos hogares para viajar. Gracias también a los que siguen cerca, en especial a Lina, Antonio Poyato, Pepe, Antonio Alba, Marta y José gracias por estar siempre ahí y por creer en mí.

A mis amigas de Antequera, las de toda la vida, gracias por crecer junto a mí. María, no tengo palabras para poder expresar tanto agradecimiento, gracias por la ayuda con las imágenes y maquetación de este trabajo. A Luisa, Carmen, Rosana, gracias por escucharme. Gracias a todas por ser mis amigas desde que mi memoria alcanza.

A mi familia, de la que estoy tremendamente orgullosa y a la que dedico este trabajo. A mi madre por su cariño y a mi padre por su capacidad de esfuerzo y fortaleza, gracias por el apoyo incondicional y por creer siempre en mí. A mi hermano, gracias por hacerme ver que con sentido del humor todo es mejor. A mi abuela, gracias por ser un ejemplo de sacrificio, esfuerzo y superación.

Finalmente, tengo que agradecer a los organismos financiadores, en especial al Instituto de Salud Carlos III por la concesión de la beca predoctoral, a la asociación de la AADEA junto con todos los clínicos de los hospitales que han participado. A los pacientes y a todas las personas que han hecho posible que se lleve a cabo este trabajo.

Indice

Abreviaturas	5
Introducción	7
1. Breve resumen de la historia de la genética	9
2. Enfermedades complejas	12
3. Estudio de las bases genéticas de las enfermedades complejas	13
4. Enfermedad de Behçet	22
5. El sistema inmune y la enfermedad de Behçet	25
6. Enfermedad de Behçet como enfermedad compleja	28
6.1 HLA en la enfermedad de Behçet	29
6.2 Regiones no HLA y la enfermedad de Behçet	33
Hipótesis, objetivos y diseño experimental	35
Publicación 1: <i>GIMAP</i> and Behçet disease: no association in the European population	43
Publicación 2: Lack of association of <i>TNFAIP3</i> and <i>JAK1</i> with Behçet's disease in the European population	45
Publicación 3: Association of <i>CCR5Δ32</i> and Behçet's disease: New data of a case-control study in the Spanish population and meta-analysis	49
Publicación 4: <i>PTPN22</i> is not associated with Behçet's disease. Study spanning the complete gene region in the Spanish population and meta-analysis of the functional variant R620W	55

Publicación 5: Variants of the <i>IFI16</i> gene affecting the levels of expression of mRNA are associated with susceptibility to Behçet disease	77
Publicación 6: Large scale analysis in Behçet disease. Identification of residues associated in the HLA Class I region and new susceptibility loci	89
Resumen de resultados y discusión	139
Objetivo 1. Replicación en población europea de asociaciones de genes/enfermedad descritas en otras poblaciones	141
Objetivo 2. Búsqueda de nuevos genes candidatos	151
Objetivo 3 Mejora el conocimiento del papel de las moléculas HLA en la enfermedad	161
Perspectivas	171
Conclusiones	175
Bibliografía	179
Anexos	195
1. Otras publicaciones del doctorando	197
2. Comunicaciones a congresos relacionadas con la tesis	199

Figuras

Figura 1. Cronología de la Genética.

Figura 2. Esquema de desequilibrio de ligamiento.

Figura 3. Plataformas de genotipado.

Figura 4. Avances en la era GWAS.

Figura 5. Esquema representativo los fundamentos de un algoritmo de imputación.

Figura 6. Representación esquemática de la distribución universal de la enfermedad de Behçet.

Figura 7. Clasificación de las enfermedades de base inmunológica.

Figura 8. Esquema de los genes de la región HLA.

Figura 9. Receptores de reconocimiento de patrones involucrados en el reconocimiento de ácidos nucleicos incluidos en el estudio.

Figura 10. Cascadas de señales en las que intervienen las proteínas IFI16 e AIM2.

Figura 11. Manhattan plot de los resultados del iChip.

Figura 12. Modelo de rutas implicadas en la fisiopatología de la enfermedad de Behçet

Tablas

Tabla 1. Cobertura del genoma de diferentes chips, calculada para poblaciones de referencia.

Tabla 2. Criterios diagnósticos del grupo internacional para el estudio de la enfermedad de Behçet.

Tabla 3. Resumen de los estudios GWAS realizados en enfermedad de Behçet.

Tabla 4. Frecuencias de los 6 SNPs de los genes *GIMAP* estudiados en diferentes poblaciones.

Tabla 5. Frecuencias de los cinco SNPs estudiados en los genes *TNFAIP3* y *JAK1* en diferentes poblaciones.

Abreviaturas

DNA	Acido desoxirribonucleico
RA	Artritis reumatoide
Arg	Arginina
RNA	Acido ribonucleico
mRNA	ARN mensajero
KIR	Receptor tipo inmunoglobulina de las células NK
Asn	Asparagina
CNV	Variación en el número de copias
BD	Enfermedad de Behçet
CD	Enfermedad de Crohn
eQTL	Loci que provocan cambios en la expresión de los genes
GWAS	Estudios de asociación de genoma completo
HSP	Proteína de choque térmico
CI	Intervalo de confianza
ICDB	Criterios diagnósticos del grupo internacional deBehçet
iChip	Immunohip
IFN	Interferon
IL	Interleucina
Kb	Kilobase
LD	Desequilibrio de ligamiento
LES	Lupus eritematoso sistémico
Lys	Lisina
MAF	Frecuencia del alelo minoritario
MB	Megabases
NK	Natural killer
OR	Odd ratio
PCA	Análisis de componentes principales
PCR	Reacción en cadena de la polimerasa
Ser	Serina
SNP	Polimorfismo de un único nucleótido
SSP	Primers de secuencia específica
Th	Células T colaboradoras
Thr	Treonina
TNF	Factor de necrosis tumoral
Trp	Triptofano
UK	Reino Unido

SSc	Esclerosis sistémica
GCA	Arteritis de células gigantes
dsDNA	ADN de cadena doble
tSNP	tagSNP
UC	Colitis ulcerosa
AS	Espondilitis anquilosante
VKH	Vogt-Koyanagi-Harada
Val	Valina

Genes

AIM2	Gen ausente en melanomas 2
CCR1	Receptor de quimioquinas 1
CCR5	Receptor de quimioquinas 5
CNTN5	Contactina 5
GIMAP	GTPasa, familia IMAP
HLA	Complejo mayor de histocompatibilidad humano
IL23R	Receptor de la interleucina 23
IFI16	Proteína inducible por interferon gamma
JAK1	Quinasa Janus 1
KIR2DL4	Receptor tipo Inmunoglobulina de las células NK con cola larga citoplasmática 4
MHC	Complejo principal de histocompatibilidad
MIC	Proteína relacionada con la cadena del MHC clase I
PTPN22	Proteína tirosina fosfatasa, no tipo receptor 22
STAT4	Transductor de señal y activador de la transcripción 4
TNFAIP3	Proteína 3 inducida por el TNF α
UBASH3B	Proteína asociada a ubiquitina que contiene dominio SH3
MEFV	Fiebre mediterránea
NOD2	Dominio de oligomerización por unión de nucleótidos

Poblaciones

ASN	Individuos del este asiático
CEU	Residentes en Utah con ancestros del norte y oeste de Europa
CHB	Chinos Han de Beijing
IBS	Población ibera en España
JPT	Japoneses de Tokio
TSI	Toscanos en Italia

Introducción

1. Breve resumen de la historia de la Genética

La revisión de la historia de cualquier área del conocimiento constituye un buen instrumento para realizar una valoración de los avances que se han producido en dicho campo. La genética es una ciencia moderna, se considera que es una ciencia del siglo XX, por lo que su historia es breve. Sin embargo, en poco más de un siglo, se ha producido una eclosión impresionante de nuevos conocimientos y avances en un sector que, a día de hoy, continúa en pleno auge expansivo.

El título de padre de la genética se le otorga a George Mendel (1822-1884) que publicó las leyes de la herencia en 1865 (aunque sus descubrimientos no alcanzaron relevancia hasta 1900). Mendel acuñó términos que son básicos en este área como alelo, dominancia y recesividad. Friedrich Miescher describió la nucleína por primera vez en 1869, pero tanto este investigador como la comunidad científica de la época seguían pensando que las moléculas responsables de transmitir la información genética de padres a hijos eran las proteínas. No fue hasta 1952 cuando se identificó, gracias a unos elegantes experimentos realizados por Alfred Hershey y Marta Chase, que el material genético estaba constituido por ácido desoxirribonucleico (DNA) aunque aún no se conocía el modo en que la información genética estaba almacenada en esta molécula. El año 1953 constituye un hito en la historia de la genética, en ese año, James Watson y Francis Crick, ayudados por estudios previos de Rosalind Franklin y Maurice Wilkins, publicaron en *Nature*, la estructura de doble hélice del DNA. Por este descubrimiento recibieron, sólo nueve años más tarde, el premio Nobel de Fisiología y Medicina, en poco más de un folio habían descrito el que supone uno de los mayores descubrimientos en la historia de la biología. El modelo que propusieron explicaba el almacenamiento de la información y permitía explicar el mecanismo de copia por el que la información se transmite a la descendencia. En 1961, Marshall Nirenberg descifró el código genético, explicando así, como la información contenida en el DNA se ve traducida en proteínas. En 1977, Frederick Sanger desarrolló una técnica para secuenciar el DNA, este método sigue utilizándose actualmente. En 1983 concurren dos acontecimientos muy

Introducción

destacables para el avance de la genética, en ese año Kary Mullis ideó la reacción en cadena de la polimerasa (PCR) técnica cuyo desarrollo, supuso un gran avance en el campo de la biología molecular, revolucionando el estudio del DNA y se identificó por primera vez el gen responsable de una enfermedad hereditaria. Otro de los grandes hitos a destacar en este resumen histórico es el lanzamiento en 1990 del Proyecto Genoma Humano, que tenía como objetivo secuenciar los 3200 millones de nucleótidos del genoma humano en 15 años. En 1999, se obtuvo la primera secuencia completa de un cromosoma humano. El proyecto llegó a término dos años antes de lo previsto y con un coste menor al presupuestado inicialmente, gracias a los extraordinarios avances en las técnicas de secuenciación que se produjeron durante ese periodo. Durante estos años, también se consiguió descifrar la secuencia completa de los genomas de diferentes organismos como *Haemophilus influenza*, *Saccharomyces cerevisiae* y *Drosophila melanogaster* entre otros. En 2002, se puso en marcha el Proyecto Internacional HapMap, cuyo objetivo era hacer un mapa del genoma humano que identificara las variaciones existentes entre individuos. El desarrollo de las nuevas plataformas de secuenciación masiva tuvo como consecuencia una caída de los precios que favoreció el desarrollo del proyecto 1000Genomas en el año 2008. Este proyecto se propuso secuenciar el genoma completo de 2500 personas procedentes de diferentes poblaciones. En este punto, ya se conocía la información relativa a la secuencia del genoma y las variaciones entre individuos, pero todavía quedan muchas interrogantes relacionadas con otros aspectos de la información genética. Para abordar algunas de las cuestiones pendientes se diseñó el proyecto ENCODE (enciclopedia de los elementos del DNA) cuya finalidad es identificar y caracterizar todos los elementos funcionales del genoma humano y cuyos primeros resultados se difundieron en 2012 (<http://unlockinglifescode.org/timeline>).

Los grandes avances producidos en tan poco tiempo, sobre todo en las dos últimas décadas, han dado lugar a una nueva era en la investigación genómica en la que se está generando un inmenso volumen de información que,

previsiblemente, tendrá un enorme impacto en muchas áreas del conocimiento científico.

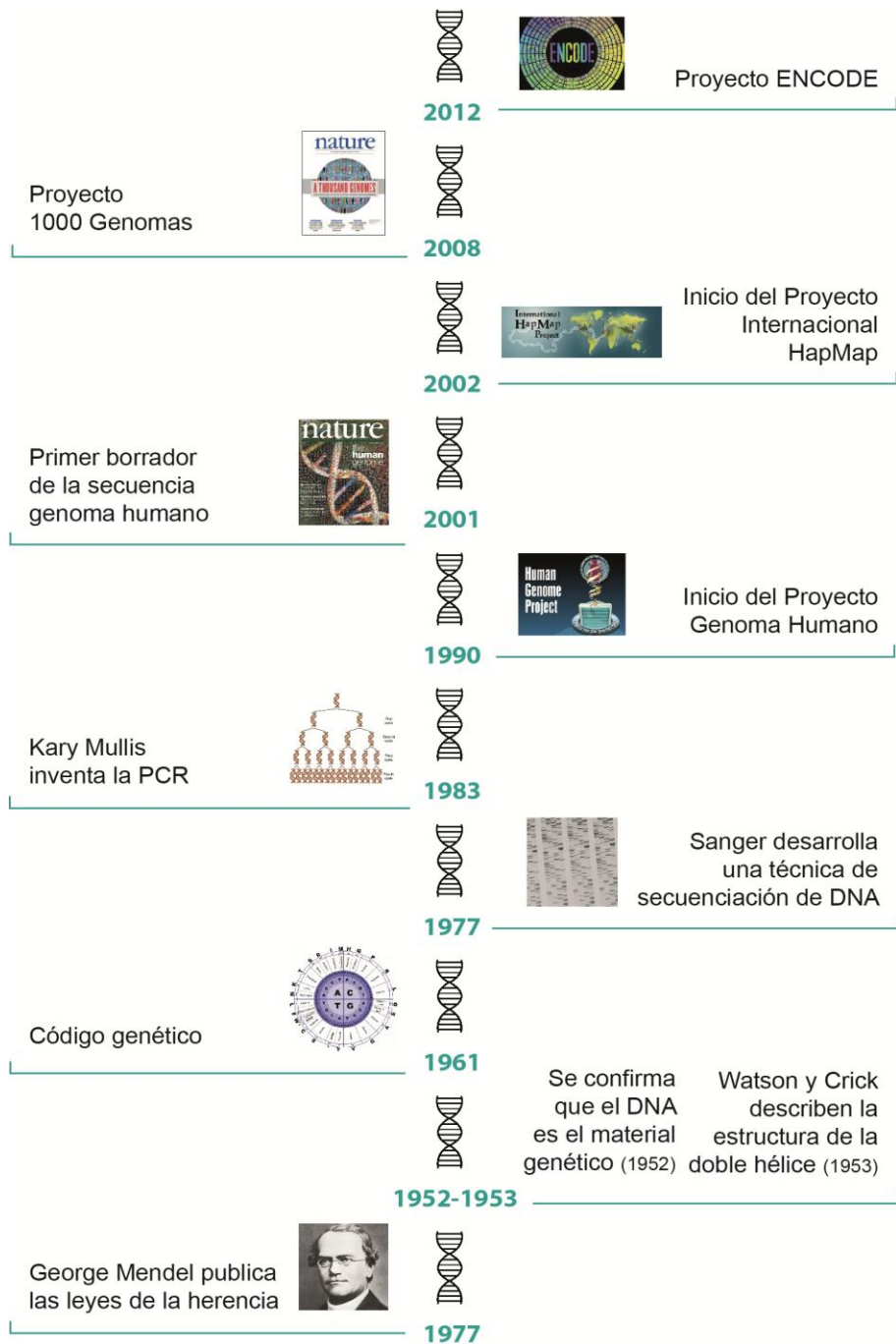


Figura 1. Cronología de la Genética. En la figura se resumen algunos de los hitos más importantes en la historia de la genética

Enfermedades complejas

La existencia de enfermedades que se transmiten a la descendencia o que son mucho más frecuentes dentro de una familia es un hecho que se conoce incluso desde antes de la publicación de los primeros estudios genéticos. Sin embargo, hasta 1983 no se realizó la primera descripción de que un gen fuera responsable de una enfermedad. Varios grupos, entre los que cabe destacar el trabajo de James Gusella, identificaron la mutación responsable de la corea de Huntington (Gusella JF *et al.*, 1983). Esta enfermedad, desde un punto de vista etiopatológico, es un ejemplo de las denominadas enfermedades clásicas o mendelianas, en las que una o varias mutaciones son las responsables directas del desarrollo de la enfermedad. Desde entonces, se han identificado los genes responsables de muchas enfermedades mendelianas mediante técnicas de ligamiento genético y clonaje posicional. Ejemplos de enfermedades mendelianas son la fibrosis quística, la anemia falciforme, la fenilcenoturia, etc. que son patologías que suelen tener graves consecuencias en cuanto a morbilidad pero que, en general, son poco frecuentes (Botstein D *et al.*, 2003). Sin embargo, otras enfermedades, más prevalentes en la población, como son, artrosis, diabetes e hipertensión arterial entre otras, son enfermedades complejas.

Se entiende por enfermedades complejas, aquellas en las que, tanto en su inicio como en su progresión, intervienen un número indeterminado de factores ambientales y genéticos. En cuanto a la base genética, son enfermedades poligénicas en las que se encuentran implicados varios, se podría decir que numerosos, genes contribuyendo cada uno de ellos al riesgo de padecer la enfermedad con un peso relativamente bajo, es decir, la penetrancia de cada gen es baja. La penetrancia es una medida de la frecuencia con la que la característica que controla el gen se manifiesta (fenotipo) en los individuos que portan un determinada variante (genotipo). Además, incluso aquellos individuos que presentan el conjunto de variantes que determinan la predisposición a la enfermedad pueden no desarrollarla, ya que, para el disparo del proceso patológico es necesario también el concurso de factores ambientales y, por otro lado, el

conjunto de variantes de riesgo puede variar de un individuo a otro y/o entre poblaciones. Es por esto, que en estas enfermedades no se habla de alelos/genes causantes, sino de alelos de riesgo, ya que, hay individuos que heredan estos alelos de susceptibilidad sin desarrollar la enfermedad e individuos que la padecen pero no los presentan (Lander ES *et al.*, 1994).

Todas estas particularidades condicionan que el abordaje del estudio genético de estas patologías sea complejo y, que a pesar de los grandes avances realizados en los últimos años, todavía quede mucho camino por recorrer hasta que se conozcan todos los mecanismos etiopatogénicos que subyacen tras estas enfermedades.

2. Estudio de las bases genéticas de enfermedades complejas

Las técnicas usadas para estudiar las enfermedades mendelianas no han sido útiles para identificar los genes que se encuentran involucrados en las enfermedades complejas. Por esta razón, los investigadores comenzaron a diseñar nuevas estrategias para abordar el estudio de estas patologías (Risch N *et al.*, 1996). Entre las estrategias más utilizadas se encuentran los estudios de asociación caso-control, si bien, los estudios familiares también han ocupado un lugar importante en la investigación de las bases genéticas de estas enfermedades. Los estudios de asociación casos-control, son estudios poblacionales en los que se compara la distribución de frecuencias de marcadores genéticos en los casos, que son los individuos a los que se le ha diagnosticado la enfermedad en cuestión, y los controles, que son individuos sanos (Lewallen S *et al.*, 1998).

Hay una gran variedad de marcadores moleculares que se pueden usar en estos estudios, entre ellos se encuentran: las repeticiones en tandem que generalmente son de pequeño tamaño (microsatélites), los polimorfismos de un único nucleótido (SNP), las variación de número de copias (CNV), etc. En los últimos años, el estudio de SNPs ha desbancado prácticamente al resto de

marcadores, de manera que, se han erigido en los marcadores por excelencia por diversos motivos.

Los SNPs son polimorfismos de un único nucleótido, en los que este cambio simple en la secuencia da lugar a distintos alelos, de los cuales, la frecuencia del alelo minoritario (MAF) tiene que ser al menos del 1%. También pueden ser inserciones o deleciones de una base en medio de una secuencia. Se calcula que en el genoma humano puede haber aproximadamente 38 millones de SNPs, lo que constituiría el 90% de todas las variaciones existentes, siendo los responsables de la mayor parte de la variabilidad genética que existe entre individuos (The 1000 Genomes Project Consortium, 2012). Otro de los motivos por lo que los SNPs se han constituido en los marcadores más utilizados ha sido el espectacular y rápido desarrollo que han experimentado las técnicas de genotipado de estos marcadores desde la aparición de la PCR, lo que ha supuesto grandes mejoras en velocidad, capacidad y accesibilidad económica en pocos años.

A pesar de estos avances tecnológicos, genotipar todos los SNPs presentes en un número elevado de individuos sigue siendo impracticable y excesivamente costoso. Además, se ha demostrado que aquellos SNPs que se encuentran cerca en el genoma muestran algún grado de desequilibrio de ligamiento (LD), es decir, las variantes tienden a heredarse juntas en más ocasiones de lo que cabría esperar por el azar, originando los denominados bloques de ligamiento (figura 2) (Lewontin RC *et al.*, 1960).

La combinación de alelos de los SNPs que componen estos bloques es lo que se conoce como haplotipos. Por tanto, para poder identificar un haplotipo no haría falta genotipar todos los SNPs que lo integran, sino un conjunto mínimo de estos a los que se les denomina tag SNPs (tSNPs) (Carlson CS *et al.*, 2004; Tenesa A *et al.*, 2006). De esta manera, se puede extraer mayor información sobre la variación genética de una región genotipando sólo los tSNPs ahorrando tiempo y dinero.

Para seleccionar los tSNPs existen programas que manejan algoritmos basados en la información existente sobre LD entre los SNPs de la población en cuestión. Esta información se encuentra recogida en las bases de datos de los

proyectos HapMap y 1000 Genomas. El valor del LD se mide utilizando 2 parámetros, D' que nos da la idea de heredabilidad de los alelos y r^2 que nos indica el coeficiente de correlación entre los alelos.

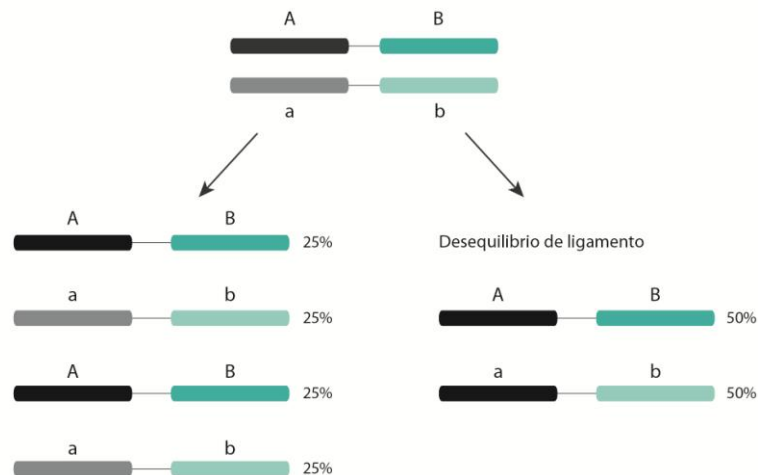


Figura 2. Esquema de desequilibrio de ligamento. La imagen representa 2 posibles situaciones extremas de herencia para 2 SNPs. Cada SNP está representado por un color, gris y verde, y las variantes están representadas por 2 tonalidades del mismo color, más oscura (A y B) y más clara (a y b). En la parte izquierda de la imagen se ejemplifica una situación en la que estos 2 SNPs se encuentran muy alejados en el genoma y no existe ningún grado de LD entre ellos. En esta situación, estos 2 SNPs se transmiten a la descendencia de manera independiente, en forma de 4 haplotipos, por combinación de las 2 variantes de los 2 SNPs, cada uno con una probabilidad de transmitirse a la descendencia del 25%. En la parte derecha se ilustra una situación de máximo LD, en esta situación, los alelos de los dos SNPs siempre se heredan juntos dos a dos. En este caso sólo se transmiten a la descendencia 2 haplotipos, cada uno con una probabilidad del 50%. En este segundo caso, basta con genotipar uno de los SNPs (tSNP) para deducir el genotipo del otro. Entre ambos casos, existe todo un gradiente de situaciones intermedias de LD.

En cuanto a las posibilidades tecnológicas, en la actualidad contamos con diferentes plataformas de genotipado con diversos fundamentos moleculares y que, según su capacidad, pueden clasificarse en plataformas de bajo, medio o alto rendimiento (figura 3). Para hacer una correcta elección de la plataforma que se

vaya a utilizar en un estudio concreto, se deberá tener en cuenta el número de individuos y de SNPs que se van a incluir en el mismo.

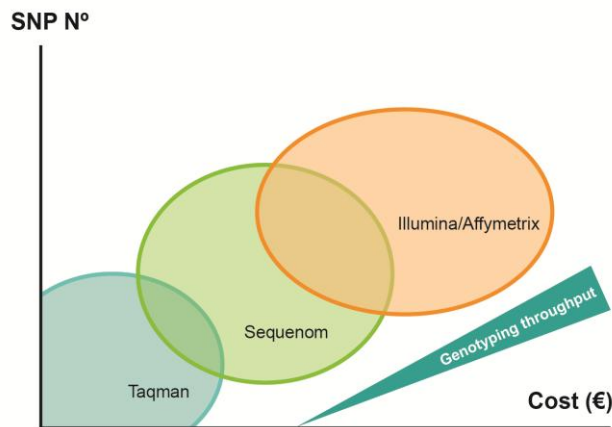


Figura 3. Plataformas de genotipado. La gráfica representa ejemplos de tecnologías de genotipado a elegir dependiendo del número de SNPs a genotipar y el coste si se considera un tamaño poblacional determinado.

En cuanto al abordaje de estudios caso/ control se pueden diferenciar en dos tipos, dependiendo de que exista una hipótesis previa para la selección de los genes, estudios de genes candidatos, o no, estudios sin hipótesis *a priori*.

La estrategia de gen candidato, por tanto, puede definirse como el estudio de la influencia de genes que son elegidos basándose en alguna hipótesis. Un gen puede ser candidato en base a diferentes motivos, puede ser elegido por un criterio funcional, es decir, porque se piense que pueda estar implicado en el desarrollo de la patología debido a la función de la molécula que codifica. La hipótesis puede basarse también en un criterio posicional, un gen que se encuentra muy próximo a otro que se haya encontrado asociado pero sin que se haya encontrado una explicación causal a dicha asociación, ya que, debido al LD existente entre variantes, la variante causal puede encontrarse lejos, incluso situada en un gen diferente a la variante que se haya encontrado asociada en un primer lugar. También puede ser buen candidato aquel gen que tenga un papel relevante en el desarrollo de una patología similar a la enfermedad de interés, bien en cuanto a la

asociación con características clínicas o a la participación en mecanismos moleculares o rutas de señalización similares (Hattersley AT *et al.*, 2005; Tabor HK *et al.*, 2002).

En cualquiera de los casos previamente descritos, la elección de los genes candidatos siempre debe ser consecuente con el conocimiento previo y tener una plausibilidad biológica. Los estudios de genes candidatos han sido profusamente utilizados y han cosechado numerosos éxitos en la investigación en esta área durante décadas. Por lo que respecta a la estrategia de estudios sin hipótesis *a priori*, éstos se basan en el rastreo de todo el genoma intentando cubrir la máxima variabilidad mediante la selección de tSNPs para cada población. Estos estudios se denominan GWAS (genome wide association study) y, a partir de los primeros, realizados en el año 2005, se han llevado a cabo en numerosas enfermedades.

El desarrollo de los GWAS ha contribuido enormemente a la mejora del conocimiento de las bases de las enfermedades complejas en la última década, como se puede apreciar en la figura 4. Además, este tipo de estudios son generadores de hipótesis y como tales también han sido muy exitosos en la investigación de estas enfermedades (Visscher PM *et al.*, 2012).

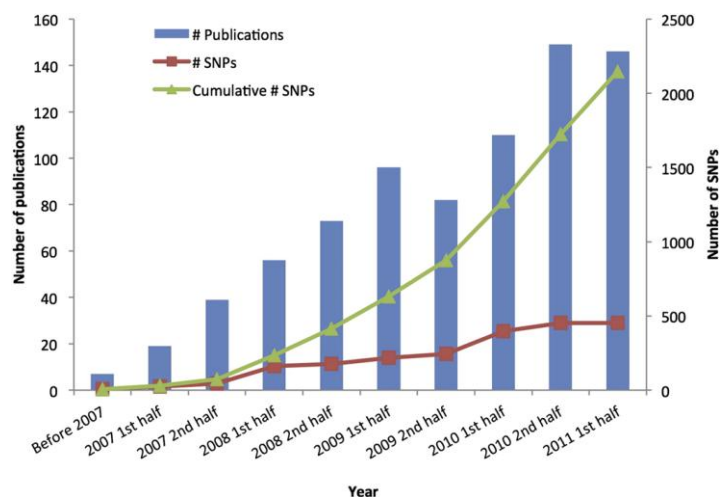


Figura 4. Avances en la era GWAS. En la gráfica se representan el número de publicaciones de GWAS y el número de SNPs que se han encontrados asociados en estos trabajos a

Introducción

lo largo del tiempo. Sólo se representan los SNPs con valores de $p < 5 \times 10^{-8}$ y se excluyen los SNPs con LD elevado ($r^2 > 0.8$). (Visscher PM *et al.* Am J Hum Genet, 2012)

Sin embargo, aún en aquellas patologías en las que se han realizado varios GWAS, una fracción de la heredabilidad sigue siendo desconocida. Se supone que esta heredabilidad perdida reside en las zonas del genoma que no están cubiertas por los chips que se han diseñado hasta el momento y/o en las variantes raras (Manolio TA *et al.*, 2009). La única técnica disponible hasta la fecha capaz de cubrir la variación de todo el genoma es la secuenciación, pero hasta el momento, es inabordable aplicarla al estudio masivo de poblaciones, sobre todo por condicionamientos económicos. En la tabla 1 se muestra el porcentaje de cobertura de variación genómica (teniendo en cuenta sólo los SNPs) de algunos de los chips comerciales que hay disponibles en el mercado usados para GWAS para población europea, asiática y africana. En el caso de la población europea la máxima cobertura que se puede alcanzar hoy día no llega al 50% (Ha NT *et al.*, 2014).

Tabla 1. Cobertura del genoma por diferentes chips, calculada para poblaciones de referencia.

	Europeos	Asiáticos	Africanos
<i><1 millón de SNPs</i>			
Axiom GW EU	0.16	0.21	0.06
Axiom GW ASI	0.17	0.21	0.06
Axiom GW SNP Array 5.0	0.15	0.18	0.05
Human OmniExpress	0.21	0.26	0.08
Human Hap550	0.18	0.22	0.06
Human Hap650Y	0.20	0.25	0.07
<i>1–2 millones de SNPs</i>			
Axiom GW CHB	0.30	0.37	0.12
Axiom GW SNP Array 6.0	0.25	0.31	0.10
Human Omni1S-8	0.22	0.29	0.10
<i>>2 millones de SNPs</i>			
Axiom GW PanAFR	0.43	0.52	0.20
Human Omni2.5-8	0.44	0.53	0.20
Human Omni2.5S-8	0.30	0.48	0.15

Los cálculos referentes a las medidas de cobertura están basados en un umbral de r^2 de 0.8 (Adaptado de Ha NT, *et al.* Eur J Hum Genet 2014)

De la mano del desarrollo tecnológico y como consecuencia del gran aumento de datos generados, se han producido también grandes avances en las técnicas de análisis y de biología computacional.

Una de las estrategias que ha recibido más atención por parte de la comunidad científica para paliar, al menos parcialmente, la mencionada pérdida de heredabilidad ha sido el desarrollo de la imputación, o en traducción más correcta, deducción, que se basa en la aplicación de algoritmos para predecir valores desconocidos a partir de datos observados. Esta estrategia se desarrolló con diferentes finalidades. Aunque la calidad de los resultados de genotipado de un chip sea adecuada, siempre existe un número relativamente elevado de polimorfismos que se pierden por no cumplir los estrictos controles de calidad que deben establecerse en estos experimentos, uno de los objetivos de los métodos de imputación es precisamente recuperar datos perdidos. Además de esta aplicación, mediante esta estrategia se consigue deducir variantes que no se encuentran incluidas en el diseño de los chips. Con ello se puede llevar a cabo un estudio en profundidad (mapeo fino) de las regiones que se han encontrado asociadas, evitando así la posibilidad de perder señales de asociación importantes. Además, el uso de estos algoritmos facilita la combinación de resultados de diferentes estudios pudiendo realizarse meta-análisis y obtener más rendimiento de los estudios ya existentes.

Se han desarrollado diferentes algoritmos, pero todos ellos comparten el mismo fundamento. Se basan en la comparación de los haplotipos de los individuos de estudio con haplotipos de los individuos existentes en las bases de datos (paneles de referencia) para inferir los datos que faltan (figura 5) (Marchini *J et al.*, 2010). De ahí la importancia de la selección de un panel de referencia adecuado para una población determinada, si se lleva a cabo la imputación con un panel de referencia inadecuado, es decir, de una población genéticamente diferente a la del estudio, se pueden obtener datos erróneos que conduzcan a conclusiones falsas.

Introducción

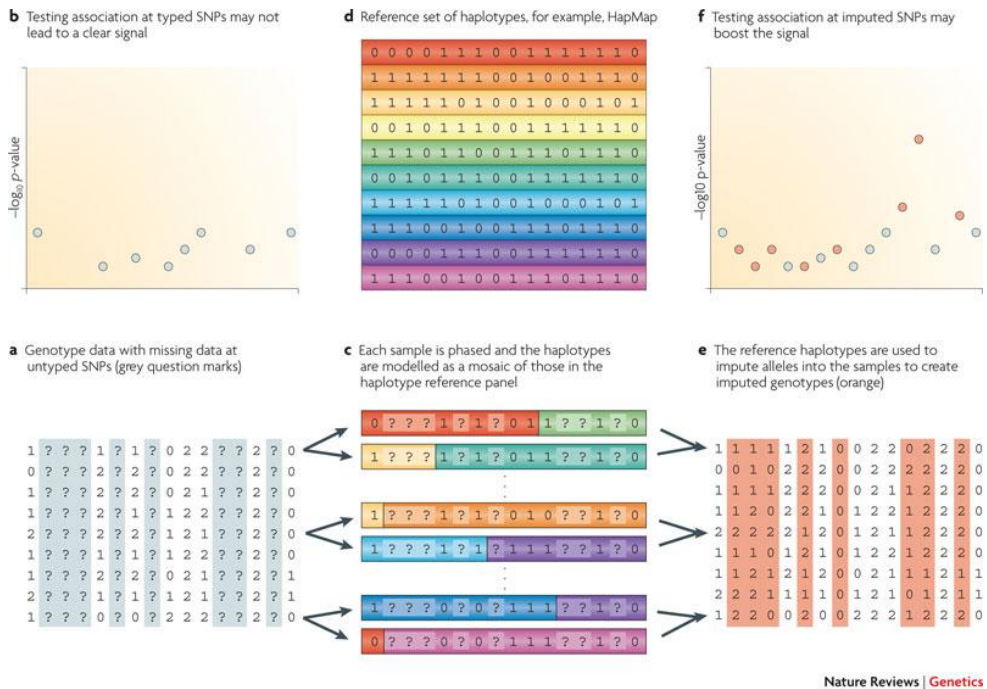


Figura 5. Esquema representativo los fundamentos de un algoritmo de imputación. En la imagen a, cada fila corresponde con un individuo y cada columna con un SNP; los genotipos están codificados como 0, 1 y 2. Los signos de interrogación corresponden con aquellos SNPs para los que no disponemos de datos. La figura b representa los logaritmos de los valores de p del estudio de asociación de los diferentes SNPs., se puede observar que no se encuentra ninguna asociación. El primer paso del algoritmo de imputación es deducir los haplotipos (procedentes de la madre y del padre) de los individuos incluidos en el estudio (c). Para lo cual, debemos seleccionar un panel de referencia adecuado a la población de estudio (d). El algoritmo compara los haplotipos y deduce la información que faltaba; los genotipos imputados se muestran en color naranja (e). Se vuelve a realizar análisis de asociación con todos los datos y se encuentra que la asociación que proviene de 1 de los SNPs que se ha imputado (f). (Marchini J, *et al* Nat Rev Genet 2010)

Otra de las causas de la pérdida de heredabilidad puede ser, como se comentaba anteriormente, la asociación con variantes raras. Este problema tiene su origen en el diseño de los chips usados para los GWAS que no suelen incluir variantes raras, ya que, se parte de la hipótesis de que las variantes genéticas que

subyacen tras las enfermedades comunes tienen que ser polimorfismos relativamente comunes en la población (Lander ES, 1996; Chakravarti A, 1999). Además, las variantes raras son más difíciles de capturar con tSNPs. Recientemente, está reforzándose la hipótesis de que variantes raras puedan ser importantes en patologías complejas (Pitchard JK, 2001). Actualmente, los estudios de genes candidatos son la mejor opción para detectar asociación de enfermedades con variantes raras. Este planteamiento impulsó el desarrollo de un chip que permitiese el genotipado a gran escala, pero cuyo diseño estuviese basado en una hipótesis previa, es decir, se diseñó un estudio de genes candidatos pero que permitiese una escala de genotipado de alto rendimiento.

El Illumina Infinium Genotyping Chip más conocido como Immunochip (iChip) se diseñó en 2009 específicamente para estudios inmunogenéticos. El objetivo de este proyecto era poder estudiar a fondo las regiones genéticas que se han encontrado asociadas a enfermedades de base inmunológica o aquellas que por su función puedan estar involucradas en dichas enfermedades. En el diseño del iChip participó un consorcio de investigadores que se basó en el conocimiento de las bases genéticas de 11 enfermedades entre las que cabe destacar artritis reumatoide (RA), psoriasis y enfermedad de Crohn (CD) entre otras. Como resultado de la información existente se diseñó un chip que incluye 196.524 polimorfismos (718 inserciones o deleciones pequeñas y 195.806 SNPs) que mapean 189 regiones del genoma. Otra característica particular de este chip es que incluye buenos marcadores de ancestría (sobre todo para población caucásica) que permiten realizar estudios de estratificación de poblaciones y evitar sesgos derivados de este problema, mejorando así el análisis. El chip también incluye un conjunto muy denso de SNPs de la región del complejo mayor de histocompatibilidad humano (HLA) y de las regiones colindantes, para poder llevar a cabo estudios de imputación de esta región. Esta característica es muy interesante ya que, esta región, como se explicará más adelante, es muy importante en las enfermedades inmuno-mediadas y su estudio es bastante complejo debido al

grado de polimorfismo tan extremadamente elevado que presenta (Cortes A *et al.*, 2011).

El estudio de las bases genéticas de diferentes enfermedades utilizando iChip ha constituido un éxito que se ha traducido en un incremento del conocimiento de estas bases en enfermedades como esclerosis sistémica (SSc), arteritis de células gigantes (GCA), etc. (Mayes MD *et al.*, 2014; Carmona FD *et al.*, 2015).

Independientemente del abordaje que se utilice, tras la identificación inicial de la asociación de un gen con una patología, se hace imprescindible realizar estudios de replicación en la misma y/o en diferentes poblaciones para validar las asociaciones descritas. Las replicaciones son muy importantes para evitar errores tipo I (falsos positivos), ya que, los estudios de asociación pueden estar influenciados por múltiples factores como pueden ser sesgos de selección, sesgos de genotipado, malas clasificaciones de grupos de estudio, factores de confusión, etc., lo que podría resultar en asociaciones falsas. Así, un resultado en el que las replicaciones en diferentes poblaciones sean consistentes indica una fuerte evidencia de verosimilitud de la asociación mientras que la falta de replicación entre poblaciones, aunque no tiene por qué implicar necesariamente lo contrario, determina la necesidad de realizar estudios complementarios para establecer o rechazar la asociación del gen con la enfermedad (Tabor HK *et al.*, 2002).

3. Enfermedad de Behçet

Aunque las manifestaciones clínicas de la enfermedad de Behçet (BD) fueron descritas por Hipócrates en el siglo V A.C. (Freigenbaum A, 1956), fue en 1937 cuando el dermatólogo turco Hulusi Behçet describió la enfermedad caracterizándola por la presencia de aftas orales, aftas genitales y uveítis (Behçet H, 1937; Behçet H, 1938; Behçet H, 1940). Se trata de una enfermedad crónica que cursa con brotes y recidivas, la edad de comienzo suele situarse en torno a los 20-30 años de edad y afecta a igual proporción entre hombres y mujeres (Yurdakul

S *et al.*, 2008). Desde el punto de vista clínico, la BD se clasifica como una vasculitis sistémica que puede afectar a vasos sanguíneos de cualquier calibre. De ahí, que el cuadro clínico que muestran los pacientes sea muy amplio y diverso. Además de las características clínicas antes mencionadas, los pacientes suelen presentar con frecuencia lesiones cutáneas, artritis, manifestaciones gastrointestinales y/o neurológicas (Mendes D *et al.*, 2009). Al no existir ningún síntoma patognomónico de la enfermedad ni tampoco ningún biomarcador específico, su diagnóstico se basa en criterios clínicos y sistemas de clasificación. Los criterios que más se utilizan actualmente son los propuestos por el grupo de estudio internacional de BD (ICBD) en 1990 (International Study group for Behcet's disease, 1990) (tabla 2).

Tabla 2. Criterios diagnósticos del grupo internacional para el estudio de la enfermedad de Behçet

CRITERIO	DESCRIPCIÓN
Ulceración oral recurrente	Aftas menores, aftas mayores o ulceración herpetiforme observadas por el clínico o por el paciente con recurrencia de al menos tres ocasiones en el periodo de un mes
Ulceración genital recurrente	Úlceras aftosas o cicatrizantes observadas por el médico o por el paciente
Lesiones oculares	Uveítis anterior, uveítis posterior o presencia de células en el vítreo en el examen con lámparas de hendidura. Vasculitis retiniana diagnosticada por un oftalmólogo
Lesiones cutáneas	Eritema nodoso observado por el médico o por el paciente, pseudofoliculitis o lesiones papulo-pustulosas, ó nódulos acneiformes observados por el médico en pacientes post-adolescentes no tratados con corticoides
Test de patergia positivo	Fenómeno de patergia positivo, evaluado por un médico a las 48 horas
DIAGNOSTICO	
Ulceración oral recurrente más otros dos criterios.	

La enfermedad también es conocida como la enfermedad de La Ruta de la Seda, debido a que la distribución geográfica de la misma coincide con la ruta que usaban los mercaderes para conectar oriente con occidente (figura 6). La prevalencia a lo largo de esta ruta no es homogénea, siendo el país con más incidencia Turquía con 421 pacientes cada 100.000 habitantes, mientras que en España la prevalencia es de 6 pacientes cada 100.000 habitantes (Cho SB *et al.*, 2012; González-Gay MA *et al.*, 2000; Graña G *et al.*, 2001). Debido a su baja prevalencia en España, se la cataloga dentro de las enfermedades raras. La relación geográfica con la prevalencia sugiere influencia de factores genéticos, ambientales o la confluencia de ambos en la predisposición a la enfermedad.

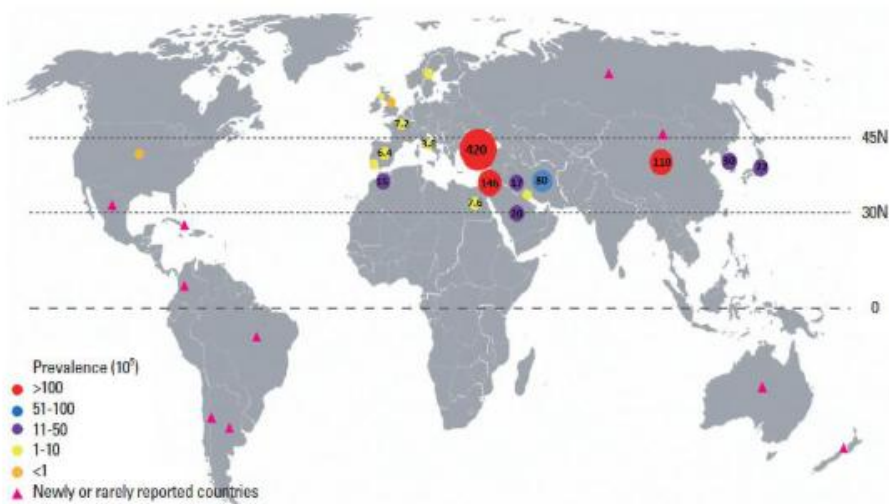


Figura 6. Representación esquemática de la distribución universal de la enfermedad de Behçet. La prevalencia de la enfermedad se representa con el tamaño y el color de los punto (Adaptada de Cho *et al.*, 2012).

La etiopatogénesis de la enfermedad sigue siendo desconocida pero algunos estudios muestran relación con agentes infecciosos. En 1997 se describió un modelo animal, en el que el 30% de los animales que eran inoculados con Herpes

virus simple tipo I presentaban síntomas de esta enfermedad, entre los que se incluían úlceras cutáneas, orales, gastrointestinales y genitales, junto con inflamación ocular y artritis (Sohn S *et al.*, 1997). Además, en determinados estudios se ha descrito que los pacientes presentan un aumento de niveles séricos de anticuerpos frente a herpes virus en comparación con controles sanos (Lee S *et al.*, 1996; Sohn S *et al.*, 1998).

Los microorganismos más estudiados en relación con BD son estreptococos concretamente se han identificado *Streptococcus sanguis*, *Streptococcus mitis* y *Streptococcus salivarius* (Kapsimali VD *et al.*, 2010) en las frecuentes infecciones nasofaríngeas y caries dentales que presentan estos pacientes. Además, algunos autores han descrito una respuesta de hipersensibilidad frente a antígenos de *Streptococcus sanguis* mediada por monocitos de sangre periférica y de la piel de los pacientes, respuesta que no se encuentra en los procedentes de individuos sanos (Kaneko F *et al.*, 2008). También se ha descrito un agravamiento de los síntomas de la enfermedad en el transcurso de infecciones dentales (Mizushima Y *et al.*, 1988). Sin embargo, aunque los microorganismos pueden estar involucrados en el origen del desajuste de los mecanismos inmunológicos que conducen a la enfermedad, no hay ninguna evidencia de que ningún microorganismo sea el causante de la misma.

4. El sistema inmune y la enfermedad de Behçet

La BD es una enfermedad de base inmunológica, clásicamente, este tipo de enfermedades se clasificaban como enfermedades autoinmunes atribuyéndose la enfermedad a fallos en el sistema inmunitario adaptativo. Más recientemente se han empezado a apreciar diferencias entre patologías. Así, la inflamación que se produce en muchas de estas enfermedades ha pasado a considerarse causa más que una mera consecuencia de las mismas. El término autoinflamación se acuñó precisamente como consecuencia de esta nueva visión de la inflamación como

Introducción

causa para englobar a aquellas enfermedades en la que el sistema inmune innato es el principal responsable de la enfermedad.

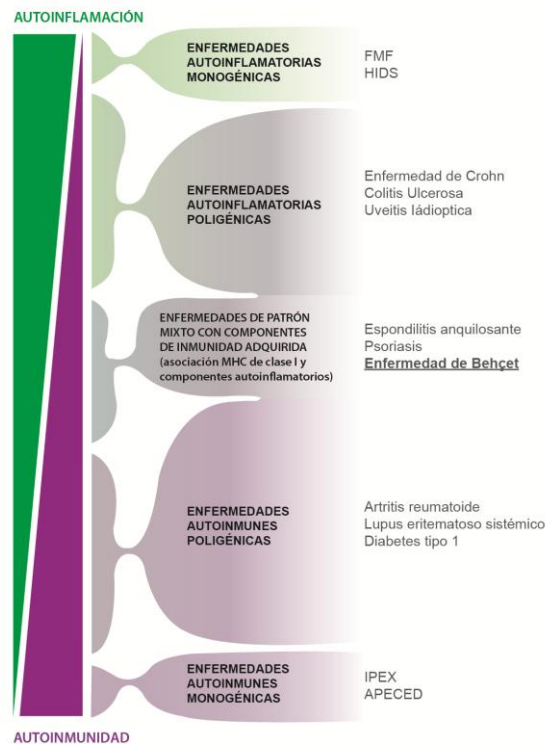


Figura 7. Clasificación de las enfermedades de base inmunológica. Existe un eje continuo en el que se clasifican las enfermedades inmuno-mediadas dependiendo del grado de implicación del sistema inmune innato (autoinflamación) y del sistema inmune adaptativo (autoinmunidad). En el centro de este eje, encontramos las enfermedades de patrón mixto, entre las que se encuentra la enfermedad de Behçet. (Adaptada de McGonagle D PLoS Med, 2006)

Hoy día, se define un eje continuo entre autoinmunidad y autoinflamación, de manera que, en cada extremo de este eje se encontrarían las raras enfermedades autoinmunes y autoinflamatorias monogénicas, en el centro se situarían las enfermedades de patrón mixto, en las que la participación de componentes de la inmunidad innata y adaptativa es similar y a ambos lados de éstas, las enfermedades poligénicas complejas de componente mayoritariamente

autoinmune como el lupus eritematoso sistémico (LES), artritis reumatoide o autoinflamatorio como la colitis ulcerosa (UC) o CD. La BD quedaría englobada precisamente en el centro, entre las enfermedades de patrón mixto, junto con otras como la espondilitis anquilosante (AS) y la psoriasis con las que comparte una serie de características que veremos más adelante (Doria A *et al.*, 2012; McGonagle MC *et al.*, 2006) (figura 7).

Se considera que los linfocitos principalmente implicados en la inmunopatogénesis de Behçet son los linfocitos T. Se han descrito numerosas alteraciones en la homeostasis de las células T ($\gamma\delta$, citotóxicas, Th1 y Th17) de estos pacientes. Sabemos que los linfocitos T $\gamma\delta$ desarrollan un papel muy importante como primera línea de defensa en la inmunidad de las mucosas y, como se ha comentado anteriormente, los pacientes con BD tienen numerosas lesiones en las mucosas, por lo que estos linfocitos han sido ampliamente estudiados en relación con la enfermedad. Aunque su papel en el desarrollo de la patología todavía no está claro, hay un incremento en cuanto a población y a activación de estas células tanto en sangre periférica como en los sitios de las lesiones mucosas (Hamazaoui K *et al.*, 1994; Freysdottir J *et al.*, 2006). Se piensa que el aumento de células activadas en la sangre de los pacientes durante la fase activa de la enfermedad se debe a la estimulación de las mismas por productos de los microorganismos presentes en las úlceras bucales (Bank I *et al.*, 2003).

Los cambios en el equilibrio Th1/Th2 se han relacionado con diversas enfermedades autoinmunes, en el caso de la BD, se ha observado un incremento de citoquinas Th1 en sangre y se ha encontrado IL12, IL18 e IFN γ en lesiones de pacientes con enfermedad en fase activa. También se han observado infiltrados de células Th1 en úlceras orales y genitales, lesiones cutáneas y gastrointestinales (Ben Ahmed M *et al.*, 2004; Dalghous AM *et al.*, 2006; Ferrante A *et al.*, 2010). Por otro lado, los monocitos activados de los pacientes producen citoquinas inflamatorias como IL1 β , IL12 y TNF α entre otras, dirigiendo una polarización de la respuesta inmune hacia la ruta Th1 (Kapsimali VD *et al.*, 2010). También se han descrito niveles elevados de IL23 e IL17 en células mononucleares de sangre

periférica de pacientes con esta enfermedad al compararlos con controles (Chi W *et al.*, 2011). La IL23 induce la producción de IL17 por los linfocitos T y esta citoquina promueve una respuesta inflamatoria mediada por neutrófilos. En las fases tempranas de las lesiones se ha observado infiltración de neutrófilos hiperactivados (Eksioglu-Demiralp E *et al.*, 2001) lo que se traduce en un incremento en los niveles de especies reactivas de oxígeno, adhesión endotelial, quimiotaxis y fagocitosis (Kobayashi M *et al.*, 2000; Carletto A *et al.*, 1997; Efthimiou J *et al.*, 1989; Neves FS *et al.*, 2009). Todos estos datos sugieren que el equilibrio Th17/Th1 es un factor importante en la regulación de la actividad de la inflamación en BD (Kim J *et al.*, 2010).

Se ha investigado el papel de varios autoanticuerpos como los anticuerpos anti-células endoteliales (Cervera R *et al.*, 1994; Aydintug AO *et al.*, 1993; Direskenelil H *et al.*, 1995), anticardiolipinas (Kawakami T *et al.*, 2009; Musabak U *et al.*, 2005) y anti-*Saccharomyces cerevisiae* (Monselise A *et al.*, 2006) en BD pero, hasta la fecha, no hay ningún dato que nos indique que ninguno de ellos sea relevante en el desarrollo de esta patología.

5. Enfermedad de Behçet como enfermedad compleja

A pesar de los avances que se han producido en el conocimiento de la enfermedad, la etiología de la BD sigue siendo en gran parte desconocida, aunque sabemos que determinados factores ambientales desencadenan la enfermedad en individuos genéticamente susceptibles. Son muchas las evidencias que sostienen que esta enfermedad tiene una base genética. Por un lado el hecho de que, como se ha comentado anteriormente, la prevalencia de la enfermedad dependa en gran medida de la localización geográfica sugiere, al menos en parte, predisposición genética. Además, hay estudios que demuestran cierta agregación familiar (Koné-Paut I *et al.*, 1999), ya que, la frecuencia de casos entre familiares de pacientes se encuentra entre 1-18%, siendo mayor en turcos (18.2%), coreanos (15.4%) y judíos (13.2%) (Kaya TI, 2012). También, aunque hay pocos estudios, se ha

descrito que la concordancia de padecer la enfermedad es mayor entre los gemelos homocigóticos que entre los dicigóticos (Masatlioglu S *et al.*, 2010).

5.1 HLA en la enfermedad de Behçet

Los genes del sistema HLA se encuentran en una región localizada en el brazo corto del cromosoma 6 que abarca unas 4 MB. Las moléculas HLA se dividen en dos grandes grupos (clase I y clase II) según su estructura, distribución celular y función.

Con respecto a la estructura, las moléculas de clase I constan de una sola cadena pesada, α , que se asocia con la cadena β 2 microglobulina; mientras que las moléculas de clase II están formadas por dos cadenas polipépticas de peso molecular similar denominadas α y β .

En cuanto a la distribución celular, las moléculas de clase I se expresan de manera constitutiva en la mayoría de las células del organismo, mientras que, en general, las moléculas de clase II se expresan de manera constitutiva sólo en células presentadoras de antígenos.

En relación con la función, las moléculas de clase I presentan péptidos a las células T CD8 y además, algunas de ellas, controlan la actividad NK a través de su unión a determinados receptores inhibidores de estas células, los denominados KIR; mientras que, las moléculas de clase II presentan péptidos a los linfocitos T CD4. Existen diferentes tipos de moléculas HLA clásicas, entendiéndose por tales las que son capaces de presentar péptidos a las células T α/β , son las moléculas A, B y C (clase I) y DR, DQ y DP (clase II). Además, existen las denominadas moléculas HLA no clásicas (como HLA-E, HLA-F, HLA-G, MIC-A, MIC-B, etc.), semejantes a las clásicas en cuanto a su estructura pero desempeñan funciones diferentes, algunas regulan la actividad de las células NK y la de los linfocitos T $\gamma\delta$.

Introducción

La región HLA es la que alberga el mayor número de asociaciones descritas para una amplia variedad de enfermedades relacionadas con el sistema inmune (de Bakker PI *et al.*, 2012)

Como en otras enfermedades de patrón mixto se ha visto que las moléculas de HLA de clase I tienen un papel importante en la enfermedad de Behçet.

Las primeras asociaciones entre *HLA* y BD fueron descritas en la década de los 70 (Ohno S *et al.*, 1973) y en ellas se relacionaba la predisposición a padecer la enfermedad con la presencia de una determinada molécula HLA-B, concretamente B51. Desde que Ohno describiera la asociación de *HLA-B51* con la BD, multitud de estudios realizados en diferentes poblaciones han obtenido resultados similares encontrando asociación entre esta molécula y la enfermedad con valores de OR bastante elevados (de Menthon M *et al.*, 2009), por lo que pudieron ser detectadas en cohortes relativamente pequeñas. Sin embargo, todavía hoy no se conocen los mecanismos involucrados en dicha asociación. Podemos afirmar que, décadas después de que se describiera esta asociación y a pesar de las múltiples y consistentes replicaciones, hay una sorprendente falta de progreso a la hora de encontrar las variantes causales de la misma y, además, siguen publicándose resultados contradictorios en algunos aspectos.

La causa principal de la aparente falta de progreso en esta cuestión se encuentra en la complejidad de la región HLA. Así, la densidad génica de esta región es muy alta, comprende sobre unos 160 genes codificantes de proteínas, con un número desproporcionadamente elevado de genes relacionados con el sistema inmunitario, entre los que se encuentran, además de los propios genes *HLA*, otros que codifican proteínas del sistema complemento, choque térmico (HSP), factor de necrosis tumoral (TNF), etc. Además, los genes se encuentran agrupados según su función, por ejemplo, en la figura 8 podemos ver como los genes que codifican las diferentes moléculas de HLA de clase I se encuentran juntos. Esta particularidad parece haberse seleccionado por presentar ventajas, por

ejemplo, en la mejora de la coordinación de la regulación de la expresión génica. Además, esta ordenación favorece el intercambio de secuencias entre genes parálogos y, por tanto, la diversificación de las diferentes familias de genes, hecho que propicia un sistema fácilmente adaptable y que sería ventajoso en términos evolutivos si tenemos en cuenta su función (Traherne JA, 2008). Se trata también de una región con un altísimo LD entre los diferentes genes, consecuencia de una tasa relativamente baja de recombinación incluso a distancias considerablemente grandes (Mmiretti MM *et al.*, 2005; The International HapMap Consortium, 2005; de Bakker PI *et al.*, 2006).

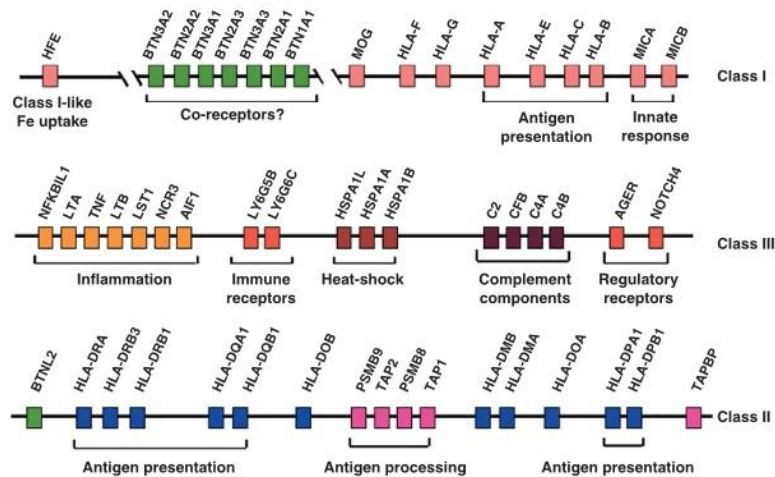


Figura 8. Esquema de los genes de la región HLA. Este esquema reducido de la región 6p21.3-22.1 en la que se observan dos tipos de agrupaciones. Por un lado encontramos genes que han sufrido duplicación, como el sistema HLA, y por otro, genes que aunque no están relacionados desde un punto de vista estructural, ya que no presentan homología en sus secuencias, si se relacionan desde el punto de vista funcional, ya que, codifican moléculas muy importantes para el sistema inmune como las moléculas del sistema del complemento. (Traherne *et al.*, 2008).

En definitiva, todos estos problemas dificultan enormemente el establecimiento definitivo de la variante causal y hay trabajos recientes que siguen poniendo en duda que dicha variante resida en el gen *HLA-B*. Como ejemplo, numerosos trabajos han estudiado la relación de la enfermedad con moléculas HLA no clásicas y dentro de ellas, la que ha tenido mayor protagonismo ha sido la

molécula MICA. El gen *MIC-A*, se encuentra situado junto a *HLA-B* y algunos estudios sugieren que la asociación de la región con la enfermedad se debe a un polimorfismo de repetición en la región transmembrana del gen *MIC-A* (Mizuki N *et al.*, 1999); sin embargo, otros estudios defienden que dicha asociación se debe al gran LD existente entre este gen y *HLA-B* (González-Escribano MF *et al.*, 1999; Yabuki K *et al.*, 1999; Wallace GR *et al.*, 1999; Mizuki N *et al.*, 2000, Salvarani C *et al.*, 2001; Cohen R *et al.*, 2002; Mizuki N *et al.*, 2002; Park SH *et al.*, 2002). Además, también se ha propuesto asociación de otros genes no HLA localizados en esta región y que son buenos candidatos, entre ellos, el más estudiado ha sido el *TNF* cuya asociación independientemente de *HLA-B* tampoco está confirmada (Zhang M *et al.*, 2013).

Centrándonos en las moléculas clásicas del sistema HLA, una de las características más reseñables es su gran variabilidad, la cual reside en su elevado grado de poligenismo y polimorfismo. En cuanto al poligenismo, ya se ha comentado que existen varios genes que codifican moléculas diferentes con la misma función. Además, cada gen es altamente polimórfico, es decir, presenta multitud de formas alélicas diferentes, algunas muy frecuentes y otras extremadamente raras y cuya frecuencia puede variar de unas poblaciones a otras. En la base de datos de inmunogenética internacional, IMGT/HLA se han documentado más de 8000 alelos. Se considera que *HLA-B* es probablemente la región más polimórfica de todo el genoma con más de 3000 variantes descritas (Robinson J *et al.*, 2015). Esta diversidad en el número de alelos, que aumentaría la probabilidad colectiva de la especie en cuanto a la capacidad de presentación de péptidos, incrementaría también las diferencias en el sistema inmune del conjunto de los individuos de la especie (Traherne JA, 2008).

A todos los problemas inherentes a las peculiaridades de esta región, hay que sumar la falta de estudios que dispongan de tamaños poblacionales adecuados, ya que no hay que olvidar que se trata de una enfermedad rara. Además, las técnicas de tipaje HLA han ido evolucionando en el tiempo, desde técnicas serológicas, menos precisas y capaces de detectar un grado de polimorfismo

bastante limitado, a técnicas de genotipado mediante biología molecular, más precisas y que permiten poner de manifiesto un mayor grado de polimorfismo. Actualmente, existen diferentes técnicas de genotipado de HLA pero todas ellas son costosas en tiempo y en dinero lo que se convierte en una limitación para llevar a cabo estudios con cohortes amplias (Erich H, 2012).

Recientemente se han publicado dos estudios a gran escala con el objetivo de profundizar en el conocimiento de las bases que sustentan la asociación de HLA con BD (Huges T *et al.*, 2013; Ombrello MJ *et al.*, 2014). Ambos trabajos se basan en los nuevos métodos de imputación y nuevos algoritmos que, además de imputar los alelos clásicos, permiten obtener información acerca de la asociación de variantes aminoácídicas en diferentes posiciones de las moléculas y de los SNPs localizados fuera de la región codificante (Jia X *et al.*, 2013). De esta manera, podemos detectar con mayor facilidad que con el tipaje clásico, si por ejemplo, la variante causal puede ser algún SNPs, que pudiera ser un eQTL y modificar, por tanto la expresión del gen, o puede ser algún aminoácido común a varias moléculas HLA. Además, la cobertura de la región es mejor, ya que incluye sitios polimórficos que podrían tener potencial informativo y que podrían estar perdiéndose si estudiamos la región únicamente a nivel de tipaje de alelos clásicos. La imputación de la región HLA, además de ser más asequible económicamente permite reanalizar los datos de chips de alto rendimiento y profundizar en el estudio de estos genes.

5.2 Regiones no HLA y enfermedad de Behçet

La asociación de la región HLA con BD sólo explica el 20-30% de la carga genética de la enfermedad (Yazici H *et al.*, 2007), lo que indica que hay genes fuera de esta región que se encuentran involucrados en la patología.

Se han realizado muchos estudios cuyo objetivo ha sido identificar los genes que son responsables del resto de la carga genética de la enfermedad pero, hasta la fecha, son muy pocas las asociaciones consistentes que se han descrito. De

nuevo, la razón principal de esta falta de éxito es el bajo poder estadístico de los estudios individuales y la falta de replicaciones en diferentes poblaciones.

Aunque los recientes análisis a gran escala han supuesto un avance en el conocimiento del componente genético de BD, está todavía lejos del de otras enfermedades inmuno-mediadas mas prevalentes, como por ejemplo de la artritis reumatoide, en las que, los estudios genéticos se han traducido en avances espectaculares en el conocimiento de la etiopatogénia. El avance en el conocimiento en el campo de enfermedades raras vendrá de la mano de proyectos colaborativos que permitan estudios con mayor poder estadístico los cuales pueden arrojar luz sobre el conocimiento de la fisiopatología de dichas enfermedades.

La tabla 3 muestra un resumen de los estudios de alto rendimiento realizados en esta enfermedad. Los resultados de estos estudios y los subsecuentes estudios de replicación, sugieren asociación de *IL-23R* y de *IL-10* con la enfermedad, pero en otros genes las asociaciones no han sido validadas porque no hay evidencia de asociación en diferentes poblaciones.

Tabla 3. Resumen de los estudios GWAS realizados en la enfermedad de Behçet.

Autor	Fei Y	Remmers EF	Mizuki N	Hou S	Lee YJ
Año publicación	2009	2010	2010	2012	2012
Población	Turcos	Turcos	Japoneses	Chinos	Coreanos
Nº Pacientes	152	1215	612	147	367
Nº Controles	172	1278	740	951	800
Chip	Affymetrix GeneChip Human Mapping 500K	Affymetrix GeneChip Human Mapping 500K	Affymetrix GeneChip Human Mapping 500K	Affymetrix genome-wide human SNP array 6.0 (1000K)	Affymetrix genome-wide human SNP array 6.0 (1000K)
SNPs analizados	Nc	311459	320438	661736	594591
Señales a GWAS level ($p < 5 \times 10^{-8}$)	Nc	Región HLA	Región HLA	Región HLA IL23R-IL12RB2 IL10	Región HLA
Señales con tendencia asociación ($p < 10^{-4}$)	LOC100129 342 KIAA1529 CPLV UBASH3B UBAC2	IL10, CPLX1, IL23R-IL12RB2		STAT4	GIMAP
Mapeo denso de genes	-	IL10, CPLX1, IL23R-IL12RB2	IL10, IL23R-IL12RB2	STAT4	GIMAP
Poblaciones de replicación	-	Turcos, Oriente Medio, griegos, UK, coreanos y japoneses	Turcos y Coreanos	Chinos	Japoneses

Nc= no consta.

$p = 5 \times 10^{-8}$ es el umbral que se usa en los estudios GWAS para considerar que existe una asociación de la variante con la enfermedad. Entre ese umbral y $p = 5 \times 10^{-4}$ se le considera zona gris, los polimorfismos aquí se encuentran en esta área, se dice que tienen tendencia a la asociación y se necesitan estudios en profundidad y estudios en otras poblaciones para establecer la verosimilitud de la asociación.

Hipótesis, objetivos y diseño experimental

Hipótesis

El esclarecimiento de las bases genéticas que subyacen a las enfermedades complejas es una cuestión de interés que está siendo objeto de investigación muy activa por parte de la comunidad científica. A pesar de que se han producido algunos avances en los últimos años, la etiopatogenia de la enfermedad de Behçet, sigue siendo bastante desconocida, si bien, se considera que es una enfermedad que se desarrolla en individuos genéticamente predispuestos por el concurso de determinados factores medioambientales. Desde un punto de vista genético se considera una enfermedad compleja en cuyo inicio y progresión deben participar numerosos genes. La mejora del conocimiento de las bases genéticas de estas enfermedades ayuda a la comprensión de las bases moleculares y, por lo tanto, de los mecanismos fisiopatológicos subyacentes a las mismas. Además, este tipo de estudios contribuyen a la descripción de nuevos marcadores que pueden ser útiles para el diagnóstico y pronóstico e incluso para postular posibles dianas terapéuticas que mejoren el tratamiento.

Objetivos

Objetivo general:

Estudiar las bases genéticas de la Enfermedad de Behçet.

Objetivos específicos:

1. Replicación en población europea de asociaciones genes/enfermedad descritas en otras poblaciones (trabajos 1, 2, 3 y 4).
2. Búsqueda de nuevos genes candidatos (trabajos 5 y 6).
3. Mejora del conocimiento del papel de las moléculas HLA en la enfermedad (trabajo 6).

Los títulos de las publicaciones relacionadas con estos objetivos son los siguientes:

1. *GIMAP* and Behçet disease: no association in the European population.
2. Lack of association of *TNFAIP3* and *JAK1* with Behçet's disease in the European population.
3. Association of *CCR5Δ32* and Behçet's disease: New data of a case-control study in the Spanish population and meta-analysis.
4. *PTPN22* is not associated with Behçet's disease. Study spanning the complete gene region in the Spanish population and meta-analysis of the functional variant R620W.
5. Variants of the *IFI16* gene affecting the levels of expression of mRNA are associated with susceptibility to Behçet disease.
6. Large scale analysis in Behçet disease. Identification of residues associated in the HLA Class I region and new susceptibility loci.

Diseño experimental

Con el propósito de alcanzar los objetivos propuestos, se planeó el siguiente diseño experimental:

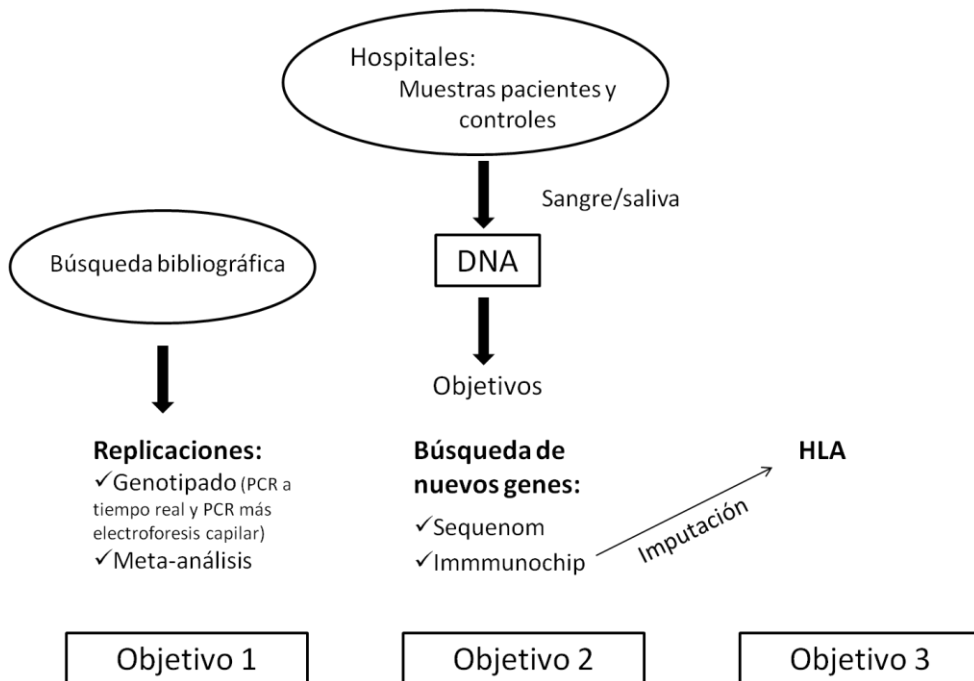
Se trata de un estudio multicéntrico en el que han participado 16 hospitales de toda España, en el cual se ha conseguido reclutar la cohorte de pacientes con BD más grande de Europa hasta la fecha. Actualmente contamos con un total de 416 pacientes, pero debido a que la recolección de muestras ha sido progresiva, la cohorte usada para cada trabajo varía dependiendo del momento en el que se realizaron los experimentos. Además, contamos con dos amplias cohortes de individuos sanos de origen español (una de aproximadamente 800 individuos, y otra de alrededor 1500 individuos), similares a la población de pacientes en términos de edad y proporción de género hombre/mujer. Las muestras biológicas de partida obtenidas tanto de los pacientes como de los controles fueron sangre y/o

saliva, a partir de estas muestras se procedió a realizar la extracción DNA y los procedimientos necesarios para alcanzar los objetivos enumerados anteriormente.

Para alcanzar el **objetivo 1**, el primer paso consistió en una revisión exhaustiva de la bibliografía referente a la genética de BD, seguida de una selección de genes potencialmente relevantes. En cuanto a los criterios de selección de los genes, por un lado escogimos aquellos que recientemente se habían descrito como asociados a la enfermedad pero solo en una población siempre que esta no fuera caucásica. Además, seleccionamos genes cuyo papel en otras enfermedades relacionadas con el sistema inmune estuviese bien establecido pero en los que los resultados publicados para BD no fueran concluyentes. Todos los polimorfismos a replicar menos uno consistían en genotipado de SNPs que se realizó con PCR a tiempo real mediante sondas TaqMan. El polimorfismo restante consistía en una delección de 32 pares de bases siendo, en este caso particular, la técnica más adecuada la PCR seguida de electroforesis capilar. En aquellos casos en los que los genes ya se habían estudiados en esta enfermedad en varias poblaciones con resultados contradictorios, además del estudio caso-control, realizamos un meta-análisis con los datos disponibles.

Para lograr el **objetivo 2**, nos planteamos dos estrategias diferentes, para cada una de las cuales elegimos la plataforma de genotipado que nos pareció más adecuada. Por un lado, realizamos un estudio de 74 tSNPs de nueve genes diferentes para el que necesitábamos una plataforma de genotipado de rendimiento medio, la plataforma escogida para este estudio fue SEQUENOM iPLEX MassARRAY que es una plataforma multiplex. Por otro lado, realizamos un estudio de alto rendimiento, para el que seleccionamos la plataforma iChip. Este chip incluye 196.524 polimorfismos, los cuales cubren un mapeo denso de 189 regiones del genoma, todas ellas relacionadas con el sistema inmunitario y abarca muy bien la gran variabilidad existente en la región HLA. Los estudios de replicación de las nuevas asociaciones identificadas se llevaron a cabo mediante PCR a tiempo real con sondas TaqMan.

Para conseguir el **objetivo 3** realizamos una imputación de los datos de genotipado de la región HLA obtenidos en el iChip, mediante el programa SNPHLA2. Este programa utiliza un algoritmo específico desarrollado para la región HLA. Mediante el uso de este algoritmo, además de deducir los alelos clásicos, también se obtiene información de las variaciones de aminoácidos en las diferentes posiciones y de los SNPs que se encuentran fuera de las regiones codificantes.



Publicaciones

GIMAP and Behçet disease: no association in the European population

Behçet disease (BD) is a multifactorial disorder with a highest prevalence among people with Mediterranean or Asian ancestry having association with HLA-B51 and other HLA and non-HLA loci.¹⁻⁴ Very recently, one genome-wide association study (GWAS) identified the *GIMAP* locus, which encoded proteins involved in development of the T and B lymphocytes, as associated with the disease in the Asian populations.⁵ To investigate whether this association is replicated in Europeans, a total of 1086 (326 patients and 760 controls) Spanish individuals of European origin were genotyped after obtaining written informed consent and approval from the local ethical committees of all the institutions involved. Patients fulfilled the 1990 International Study Group classification criteria for Behçet's disease,⁶ and the group was composed of 44.5% men with the following clinical features: 100% had oral ulcers, 63.6% genital ulcers, 58.7% uveitis, 48% arthritis, 20.2% vascular, 21.9% neurological and 19.2% gastrointestinal involvement. Regarding HLA association, in addition to B*51 ($p < 10^{-7}$, OR=4.11, 95% CI 2.79 to 6.06), B*57 was found to be risk factor in this cohort with an OR similar to that provided by B*51 ($p = 0.00008$, OR=3.70 95% CI 1.78 to 7.83).⁷ The a priori statistical power (SP) values were calculated using the CaTS Power Calculator (<http://www.sph.umich.edu/csg/abecasis/CaTS/>), and they were greater than 85% for the allelic as well as for the dominant models at the 5% significance level with the ORs previously described in the Korean population⁵ and the minor allele frequencies described for these SNPs in the CEU (Utah residents with ancestry from northern and western Europe) population.⁸ Frequencies of the risk alleles of the single nucleotide polymorphisms (SNPs) within *GIMAP4* and *GIMAP2*: rs1916012A, rs1522596A, rs1608157C, rs10266069A and rs10256482T were higher in the Spanish (0.577, 0.583, 0.575, 0.489 and 0.584, respectively) than in the Korean (0.484, 0.484, 0.484, 0.411 and 0.425, respectively) and the Japanese (0.410, 0.410, 0.408, 0.403 and 0.403, respectively) healthy populations. The frequency of the rs2286900A, the risk allele in the *GIMAP1* region, was lower in our population (0.085) than in the Korean (0.141) and in

the Japanese (0.196) populations. The SP calculated according to the allelic frequencies in our controls were also greater than 85% to detect an association similar to that described in Koreans for the allelic as well as for the dominant models, at the 5% significance level. No statistically significant differences were observed among the patient and the control groups either in the allelic or in the dominant models (table 1). Therefore, in spite of an adequate SP, the association between the *GIMAP* and BD was not replicated in our cohort. In agreement with our results, no association of this region with BD has been reported in Europeans. There are differences among populations in the set of genes influencing a specific pathology,⁹ for example, Lee *et al*⁵ reported not having found an association of this disease with the *IL23R-IL12RB2* and *IL-10* regions which have been associated with BD in several studies and in different populations.²⁻⁴ Discrepancies could be due, among other reasons, to the different features of the cohorts mainly when the primary association is with clinical manifestations. Unfortunately, we have no power for analysis according to clinical findings or HLA groups and, therefore, we cannot discard association with specific subgroup of the disease. In conclusion, the association described in Asian populations between the *GIMAP* region and Behçet's disease was not replicated in Europeans, but influence of this region in clinical features cannot be ruled out. DNA was obtained from peripheral blood or saliva using previously described standard methods. All the SNPs were genotyped using TaqMan SNP Genotyping Assays (Applied Biosystems, Barcelona, Spain) in a 7500 Real Time PCR System (Applied Biosystems) following the manufacturer's instructions. About 10% of the samples were genotyped twice to verify the genotyping consistency, showing 99% identical genotypes, and the genotype success rate was higher than 95% for all the SNPs.

The χ^2 test was used to test Hardy-Weinberg equilibrium (HWE), and to compare allelic and genotype distributions in control and patient groups. OR and 95% CI were calculated according to the Woolf method. The study population was found to be in the HWE for all the polymorphisms analysed ($p > 0.05$). None of the p values were lesser than 0.05, the value that we considered statistically significant.

The Linux software PLINK V.1.07 (<http://pngu.mgh.harvard.edu/purcell/plink/>) was used to carry out all the statistical analyses.¹⁰

Table 1 Distribution of the *GIMAP* alleles of the 5 SNPs included in the study among Spanish Behçet disease (BD) patients* and healthy controls†

		Allele frequencies		Allelic model		Dominant model	
Alleles		Cases n=326	Controls n=760	p Value	OR (95% CI)	p Value	OR (95% CI)
GIMAP 4							
rs1916012	A	0.571	0.577	0.802	0.976 (0.808 to 1.179)	0.524	0.895 (0.636 to 1.26)
rs1522596	A	0.581	0.583	0.939	0.993 (0.822 to 1.2)	0.799	0.956 (0.674 to 1.355)
rs1608157	C	0.564	0.575	0.654	0.958 (0.792 to 1.158)	0.509	0.891 (0.634 to 1.254)
GIMAP 2							
rs10266069	A	0.506	0.489	0.468	1.071 (0.889 to 1.29)	0.371	1.15 (0.847 to 1.56)
rs10256482	T	0.600	0.584	0.510	1.07 (0.880 to 1.293)	0.574	1.112 (0.769 to 1.608)
GIMAP 1							
rs2286900	A	0.099	0.085	0.293	1.184 (0.864 to 1.624)	0.19	1.254 (0.894 to 1.761)

*Patients fulfilled the 1990 International Study Group classification criteria for BD.⁶

†Bone marrow and blood donors were included as healthy controls.

Lourdes Ortiz-Fernández,¹ Marta Conde-Jaldón,¹
 José Raul García-Lozano,¹ Marco Antonio Montes-Cano,¹
 Norberto Ortego-Centeno,² María Jesús Castillo-Palma,³
 Gerard Espinosa,⁴ Genaro Graña-Gil,⁵ Juan Sánchez-Bursón,⁶
 Miguel Angel González-Gay,⁷ Ana Celia Barnosí-Marín,⁸ Roser Solans,⁹
 Patricia Fanlo,¹⁰ Mónica Rodríguez Carballeira,¹¹ Teresa Camps,¹²
 Santos Castañeda,¹³ Javier Martín,¹⁴ María Francisca González-Escribano¹

¹Servicio de Inmunología, IBIS, H.U. Virgen del Rocío/CSIC/Universidad de Sevilla, Sevilla, Spain

²Servicio de Medicina Interna, H. Clínico San Cecilio, Granada, Spain

³Servicio de Medicina Interna, H. U. Virgen del Rocío, Sevilla, Spain

⁴Servicio de Enfermedades Autoinmunes, H. Clinic, Barcelona, Spain

⁵Servicio de Reumatología, CHU A Coruña, A Coruña, Spain

⁶Servicio de Reumatología, H. de Valme, Sevilla, Spain

⁷Servicio de Reumatología, H. Marques de Valdecilla, Santander, Spain

⁸Servicio de Medicina Interna, H. Torrecárdenas, Almería, Spain

⁹Servicio de Medicina Interna, H. Vall d'Hebron, Barcelona, Spain

¹⁰Servicio de Medicina Interna, H. Virgen del Camino, Pamplona, Spain

¹¹Medicina Interna Unitat de Malalties Autoimmunes i Sistèmiques Hospital Universitari Mútua Terrassa, Terrassa, Spain

¹²Servicio de Medicina Interna, H. Carlos Haya, Málaga, Spain

¹³Servicio de Reumatología, H. de la Princesa, Madrid, Spain

¹⁴Instituto de Parasitología y Biomedicina López Neyra, Granada, Spain

Correspondence to Dr María Francisca González-Escribano, Servicio de Inmunología. HU Virgen del Rocío, Sevilla 41013, Spain; mariaf.gonzalez.sspa@juntadeandalucia.es

Acknowledgements The authors would like to thank the Asociación Andaluza de Enfermedades Autoinmunes (AADEA) and all patients and donors enrolled in the present study for their cooperation.

Contributors MAM-C, JRG-L and MFG-E contributed to the conception and design of the research, analysis and interpretation of the data, draft and revision of the manuscript. LO-F and MC-J performed the experiments and analysed the data. The rest of the authors: NO-C, MJC-P, GE, GG-G, MAG-G, ACB-M, RS, PF, MRC, TC, SC, JS-B and JM contributed with samples and patient information. All authors read and approved the final version.

Funding This work was supported by Fondo de Investigaciones Sanitarias (10/1701), Fondos FEDER, Plan Andaluz de Investigación (CTS-0197 and CTS-180), Red Enfermedades Inflamatorias y Reumáticas and Consejería de Salud de la Junta de Andalucía (0260/08). LOF is the recipient of a fellowship (FI11/00547).

Competing interests None.

Patient consent Obtained.

Ethics approval The local ethical committees of all the institutions involved.

Provenance and peer review Not commissioned; externally peer reviewed.

To cite Ortiz-Fernández L, Conde-Jaldón M, García-Lozano JR, et al. *Ann Rheum Dis* Published Online First: [please include Day Month Year] doi:10.1136/annrheumdis-2013-205156

Received 26 December 2013

Revised 20 February 2014

Accepted 1 March 2014

Ann Rheum Dis 2014;0:1–2. doi:10.1136/annrheumdis-2013-205156

REFERENCES

- Mendes D, Correia M, Barbedo M, et al. Behçet's disease: a contemporary review. *J Autoimmun* 2009;32:178–88.
- Hughes T, Coit P, Adler A, et al. Identification of multiple independent susceptibility loci in the HLA region in Behçet's disease. *Nat Genet* 2013;45:319–24.
- Remmers EF, Cosan F, Kirino Y, et al. Genome-wide association study identifies variants in the MHC class I, IL10, and IL23R-IL12RB2 regions associated with Behçet's disease. *Nat Genet* 2010;42:698–702.
- Mizuki N, Meguro A, Ota M, et al. Genome-wide association studies identify IL23R-IL12RB2 and IL10 as Behçet's disease susceptibility loci. *Nat Genet* 2010;42:703–6.
- Lee YJ, Horie Y, Wallace GR, et al. Genome-wide association study identifies GIMAP as a novel susceptibility locus for Behçet's disease. *Ann Rheum Dis* 2013;72:1510–16.
- International Study group for Behçet's disease. Criteria for diagnosis of Behçet's disease. *Lancet* 1990;335:1078–80.
- Montes-Cano M, Conde-Jaldón M, García-Lozano J, et al. HLA and non-HLA genes in Behçet's disease: a multicentric study in the Spanish population. *Arthritis Res Ther* 2013;15:R145.
- HapMap Database [http://hapmap.ncbi.nlm.nih.gov/cgi-perl/gbrowse/hapmap24_B36/]
- Rosenberg NA, Huang L, Jewett EM, et al. Genome-wide association studies in diverse populations. *Nat Rev Genet* 2010;11:356–66.
- Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:598559–75.

Lack of association of *TNFAIP3* and *JAK1* with Behçet's disease in the European population

L. Ortiz-Fernández¹, J.-R. García-Lozano¹, M.-A. Montes-Cano¹, M. Conde-Jaldón¹, N. Ortego-Centeno², F.-J. García-Hernández³, G. Espinosa⁴, G. Graña-Gil⁵, J. Sánchez-Bursón⁶, M.R. Juliá⁷, R. Blanco⁸, A.-C. Barnosi-Marín⁹, R. Solans¹⁰, P. Fanlo¹¹, M. Rodríguez Carballeira¹², T. Camps¹³, S. Castañeda¹⁴, J. Martín¹⁵, M.-F. González-Escribano¹

Lourdes Ortiz-Fernández, PhD*
José-Raúl García-Lozano, PhD, MD
Marco-Antonio Montes-Cano, PhD
Marta Conde-Jaldón, PhD
Norberto Ortego-Centeno, MD
Francisco-José García-Hernández, MD
Gerard Espinosa, MD
Genaro Graña-Gil, MD
Juan Sánchez-Bursón, MD
María Rosa Juliá, PhD
Ricardo Blanco, MD
Ana-Celia Barnosi-Marín, MD
Roser Solans, MD
Patricia Fanlo, MD
Mónica Rodríguez Carballeira, MD
Teresa Camps, MD
Santos Castañeda, MD
Javier Martín, MD
María-Francisca González-Escribano, PhD

This work should be attributed to:
Servicio de Inmunología, IBiS, Hospital
Universitario Virgen del Rocío/CSIC/
Universidad de Sevilla, Sevilla, Spain.

*List of authors' affiliations on page 4.

Please address correspondence
and reprint requests to:

María Francisca González-Escribano,
Servicio de Inmunología,
H.U. Virgen del Rocío,
41013 Sevilla. Spain.

E-mail:

mariaf.gonzalez.sspa@juntadeandalucia.es

Received on October 22, 2014; accepted
in revised form on January 26, 2015.

© Copyright CLINICAL AND

EXPERIMENTAL RHEUMATOLOGY 2015.

Key words: Behçet's disease, *JAK1*,
TNFAIP3, genetic contribution

Funding: this work was supported by
Fondo de Investigaciones Sanitarias
(10/1701 and 13/01118), Fondos FEDER,
Plan Andaluz de Investigación (CTS-0197),
Red Enfermedades Inflamatorias y
Reumáticas (RD08/0075/0013) and
Consejería de Salud de la Junta de
Andalucía (PI411/10).

L. Ortiz-Fernández is the recipient of a
fellowship (FI11/00547).

Competing interests: none declared.

ABSTRACT

Objective. Behçet's disease (BD) is an immune-mediated and complex disease which has been associated with HLA class I molecules although other genes such as IL23R and IL10 have also been involved in the susceptibility to BD. Recently, an association of variants of the *JAK1* and *TNFAIP3* genes with the disease has been reported in the Chinese Han population. The aim of the present work was to assess whether the association described in Asian populations is replicated in Europeans.

Methods. This study includes a total of 1155 Spanish subjects of European origin (372 BD and 783 unrelated healthy individuals). Patients were recruited from different hospitals and controls were collected in the same geographic regions and they matched with patients in age and gender. A total of five SNPs, two in the *JAK1* gene: rs2780815 and rs310241 and the other three in the *TNFAIP3*: rs10499194, rs9494885 and rs610604, were included in this study. The genotyping of these SNPs was performed using a real time PCR system (TaqMan® SNP Genotyping Assays).

Results. No statistically significant differences were found when the patient and control groups were compared. The distribution of the risk alleles was similar in patients with and without eye manifestations and in patients with and without HLA-B*51.

Conclusion. The association of variants of the genes *JAK1* and the *TNFAIP3* with BD which has been described in the Chinese population was not replicated in Europeans.

Introduction

Behçet's disease (BD) is an immune-mediated and complex disease in which

certain environmental factors such as infectious agents are the triggers of the disease in genetically predisposed individuals (1). This rare disease is most common along the old route named "Silk Road", stretching from China to the Mediterranean area (2). This particular geographical distribution in addition to the familial aggregation and association with HLA class I molecules (specifically with HLA-B51) are evidences supporting genetic contribution into the pathogenesis of the disease (3, 4). The contribution of the HLA region to the genetic component has been estimated to be approximately 20% (5). Some genes such as IL23R, IL10 and others have been related with BD in different populations (6-10), whereas in other cases the association seems to be limited to one ethnic group or specific population (11-14). The complexity of this vasculitis which is clinically characterized by the presence of oral and genital ulcerations but whose manifestations are diverse and they can spread to eyes, joints, digestive system, etc. (1) is an additional difficulty to establish whether a gene is related with the disease itself or with a specific subgroup. The *JAK1* (Janus kinase 1) gene is located in 1p32.3-31.3, this gene encodes a large and widely expressed membrane-associated phosphoprotein with protein-tyrosine kinase activity, which is involved in the signal transduction pathways of interferon-alpha/beta and gamma (15). The *TNFAIP3* (Tumour necrosis factor alpha-induced protein 3) gene is located at 6q23, this gene encodes an ubiquitin-editing enzyme (also known as A20) which is involved in NF-kappa-B responses and as a consequence, it is a key regulator in different immune and inflammatory responses.

es (16). Therefore, from a functional perspective, these two genes are good candidates for association with immuno-mediated diseases; in fact, they have been related with different autoimmune diseases (17-19). Recently, an association of variants of the genes *JAK1* and *TNFAIP3* with BD has been reported in the Chinese Han population (20, 21) in patient cohorts with 100% uveitis. Nevertheless, the association of these genes with BD has been not yet investigated in Europeans neither in BD patients without ocular involvement.

The aim of the present work was to assess whether the association of *JAK1* and *TNFAIP3* described in the Chinese Han population is replicated in Europeans.

Material and methods

Subject of study

The study includes 372 BD patients (161 males and 211 females) with a mean age at onset (years) SD of 48.22 ± 12.19 who fulfilled the criteria of 1990 of the International Study Group for the classification of Behçet's disease (22) and 783 unrelated healthy individuals recruited in the same geographic regions who matched in age and gender with BD patients. All the subjects were Spanish from European origin and they were recruited from different Spanish hospitals. The study was approved by the local ethics committees of all the participant hospitals, A Coruña (CHU A Coruña), Almería (H. Torredacárdenas), Barcelona (H. Clinic, Vall d'Hebron and Mútua Terrassa), Granada (H. Clínico San Cecilio), Madrid (H. de la Princesa), Málaga (H. Carlos Haya), Palma de Mallorca (H. Universitari Son Espases), Pamplona (H. Virgen del Camino), Santander (H. Marques de Valdecilla) and Sevilla (H. Virgen del Rocío and H. Virgen de Valme) and a written informed consent was obtained from all participants. Clinical features of the patient group were: 100% had oral ulcers, 66% genital ulcers, 59% uveitis, 51.6% arthritis, and 22% vascular, 23% neurological and 20% gastrointestinal involvement.

DNA extraction

Peripheral blood or saliva were used as the starting material. Genomic DNA

was extracted using the QIAamp DNA Mini Kit (Quiagen, Barcelona, Spain) according to the manufacturer's recommendations and stored at -20°C until use. The purity of DNA was determined using NanoDrop spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA). Only those DNA samples having a 260/280 ratio between 1.7 and 2.0 and a final concentration of 10-20 ng/ μl were genotyped.

SNP selection and genotyping

Regarding the *JAK1* gene, we selected two out of the three SNPs whose association to BD has been reported in the Chinese Han population (20). The minor allele frequency (MAF) values of the two selected SNPs, rs2780815 and rs310241, in Europeans are 0.47 and 0.24, respectively. The other SNP associated in the Chinese population, rs3790532, was excluded because it is not polymorphic in Europeans (MAF is zero) and, in addition, it is in a complete linkage disequilibrium (LD) with the rs310241 ($r^2=1$; $D'=1$) in the Asian population (<http://www.1000genomes.org>). Concerning the *TNFAIP3* gene, three SNPs were found associated with BD in the Chinese Han population, rs10499194, rs7753873 and rs9494885 (21). From these, we choose rs10499194 and rs9494885 (MAF 0.26 and 0.08, respectively in Europeans). The last is in a strong LD with rs7753873 ($r^2=1$; $D'=1$) in Asian and European populations. Another SNP in the *TNFAIP3* was also included in the present study, rs610604. This SNP was genotyped in the Chinese cohort and statistically significant differences in the distribution of frequencies in patients and controls were not found. Nevertheless, we decided to include this SNP because the MAF in Europeans is higher than in Asians (0.33 vs. 0.09) and, as a consequence, the statistical power for this SNP is higher in our population assuming that the rest of the conditions were similar. Genotyping of these SNPs was performed using TaqMan® SNP Genotyping Assays (Applied Biosystems, Barcelona, Spain) in a LightCycler 480 (Roche, Barcelona, Spain). About 10% of the samples were genotyped twice to verify the genotyping consistency,

showing identical genotypes in 99% of the cases. Patients and controls had been previously genotyped in HLA-B using PCR-SSOP Luminex method with LABType SSO (One Lambda Inc., Canoga Park, CA), following the manufacturer's instructions (23).

Statistical analysis

The Linux software PLINK v.1.07 (<http://pngu.mgh.harvard.edu/purcell/plink/>) was used to carry out all the statistical analyses (24). The χ^2 test was used to test Hardy-Weinberg equilibrium (HWE) and to compare genotype distribution in control and patient groups. Univariate analysis of the distribution of the frequencies of genotypes and alleles were performed in 2x2 contingency tables using the χ^2 or Fisher exact tests, when appropriate. The p -values < 0.05 were considered statistically significant and the odds ratios (ORs) and 95% CIs were calculated according to the Woolf method. The statistical powers to detect associations with ORs ranging from 1.1 to 1.6 were calculated in the allelic model taking into account the frequency of each variant in our population with the Statistical Power test of the Research Tool Kit of the DSS Research Software <https://www.dssresearch.com/KnowledgeCenter/toolkitcalculators/statisticalpowercalculators.aspx>.

Results

In the cohort included in our study, the successful rate of genotyping was 95%, the study population was found to be in the Hardy-Weinberg equilibrium ($p > 0.05$) and distribution of the variants in *JAK1* and *TNFAIP3* in the cohorts from different hospitals was not significantly different. Considering the allelic model, the statistical power values for the 2 SNPs of *JAK1* (rs2780815 and rs310241) and the 3 SNPs of *TNFAIP3* (rs10499194, rs610604 and rs9494885) were greater than 80% to detect an OR=1.3 (Table I).

The distribution of the genotypes and MAFs of the five SNPs in the patient and control groups are displayed in Table II. The MAFs of the five SNPs in our control population were not significantly different to those described in 1000 Genome Project database for

the European population. According to these data, no statistically significant differences were found when the patient and control groups were compared. The Table II also displays that the distribution of the MAFs of these SNPs in the groups of patient with ocular involvement and HLA-B*51 positive are not different to those observed in the whole group.

Discussion

The conclusion of the present study is that the association recently described in the Asian populations between the two genes included in this study, *JAK1* and *TNFAIP3*, and BD is not replicated in Europeans. The frequencies of the five SNPs included in our study were not significant different in the Spanish population when patients and the controls were compared. The consolidation of the findings of genetic association studies in complex diseases requires replication in different cohorts to avoid different types of bias. It is necessary to discard a false negative result in our study (type II error) in order to conclude that these genes are not associated in the Spanish population, therefore it is important to demonstrate an adequate statistical power. In this sense, it is unlikely that the lack of association detected in our study was due to a false negative result because it had a statistical power greater than 80%, to detect similar ORs to those described in the Chinese Han population at a 5% significance level. Discarded type II errors, discrepancies may be originated by false positive results in the original study (type I errors). In this sense, we note that the findings in Chinese population were published in two different reports with a design of cases and controls and using a candidate gene approach but only in the case of the *JAK1* gene the study included two cohorts (20), in the case of the *TNFAIP3* gene, the results were not confirmed in a second cohort (21). Even in the case of the *JAK1* the discovery and replication cohorts were from the same population (20). The lack of association is also supported by the genotyping studies performed in BD using high throughput platforms. Although these studies

Table I. Statistical power at the 5% significance level, taking into account the allelic frequencies in our population for a set of values of ORs from 1.1 to 1.6. The values for all the SNPs included in this study are shown in this table.

	Statistical power for each OR value					
	1.1	1.2	1.3	1.4	1.5	1.6
<i>TNFAIP3</i>						
rs10499194	26	61	87	97	99	100
rs610604	27	62	88	98	99	100
rs9494885	18	40	65	84	93	98
<i>JAK1</i>						
rs2780815	28	65	90	98	99	100
rs310241	25	59	86	97	99	100

Table II. Minor allele frequencies (MAFs) of the five SNPs included in this study in Spanish BD patients and healthy controls. MAFs of these five SNPs in the subgroup of patients with ocular involvement and HLA-B*51 positive are also displayed.

	Allele	BD Patients			Controls	
		Whole group n=372	Ocular involvement n=219	B*51 + n=153*	Whole group n=783*	B*51 + n=48*
<i>JAK1</i>						
rs2780815	G	0.49	0.47	0.46	0.47	0.48
rs310241	G	0.27	0.28	0.27	0.28	0.29
<i>TNFAIP3</i>						
rs10499194	T	0.32	0.34	0.37	0.30	0.35
rs61064	G	0.36	0.34	0.35	0.33	0.31
rs9494885	C	0.11	0.11	0.13	0.12	0.11

*Not all the individuals were genotyped in HLA-B. A total of 356 patients and 357 controls were genotyped. Comparisons made were patients vs. controls; patients with ocular involvement vs. patients without this manifestation; patients with ocular involvement vs. controls; patients B51 positive vs. patients B51 negative; patients B51 positive vs. controls B51 positive. The p-values were greater than 0.05 in all the cases.

can not be considered replication studies because data about whether these gene regions were excluded from the final analysis after applying the quality criteria are not available, none of these studies has identified *TNFAIP3* or *JAK1* as risk factors for BD in Europeans (6,7, 25-27). Other factor to take into account is that the genotyping method may influence the ratio of typing errors. Our study was performed using real time PCR whereas in the two original studies, typing were performed using the IPLEX system in the case of *JAK1* (20) and PCR followed by restriction fragment length polymorphism analysis in the case of *TNFAIP3* (21). Although, theoretically, the number of errors attributable to the method would be the same in the patient and control cohorts if the same genotyping method is used; actually, the percentage of error in each group (patients and

controls) may be different if the quality of the DNA is different. Moreover, when the primary association is mainly related with clinical manifestations, the discrepancies could be due to the different features of the cohorts. Regarding this point, in the Chinese cohorts, 100% of the patients have eye manifestations, therefore association with the disease is indistinguishable from association with ocular involvement in these studies. In our cohort 59% of patients have ocular manifestations but our results do not support association with this clinical feature because the distribution of the MAFs of these SNPs in patients with and without ocular manifestations was the same. Some times the association with a specific factor depends on the presence of other genetic factor, if only because this factor defines a more homogeneous group of patients. For this reason, our cohort

was stratified according to the presence of HLA-B*51 but no differences in the distribution of the MAFs were observed. Finally, the set of genes influencing a specific pathology may have differences among populations. There are multiple examples of this fact, for example, the regions *IL23R-IL12RB2* and *IL-10* have been associated with BD in several studies (6, 7). However, Lee *et al.* reported no association (11). In conclusion, the association described in the Chinese population between the *JAK1* and the *TNFAIP3* genes and Behçet's disease was not replicated in Europeans.

Acknowledgments

The authors would like to thank the Asociación Andaluza de Enfermedades Autoinmunes (AADEA) and all patients and donors enrolled in the present study for their cooperation.

Authors' affiliations

¹Servicio de Inmunología, IBI, Hospital Universitario Virgen del Rocío/CSIC/ Universidad de Sevilla, Sevilla;

²Servicio de Medicina Interna, Hospital Clínico San Cecilio, Granada;

³Servicio de Medicina Interna, Hospital Universitario Virgen del Rocío, Sevilla;

⁴Servicio de Enfermedades Autoinmunes, Hospital Clinic, Barcelona;

⁵Servicio de Reumatología, Complejo Hospitalario Universitario, A Coruña;

⁶Servicio de Reumatología, Hospital Universitario de Valme, Sevilla;

⁷Servei d'Immunologia. Hospital Universitari Son Espases, Palma de Mallorca;

⁸Servicio de Reumatología, Hospital Marqués de Valdecilla, Santander;

⁹Servicio de Medicina Interna, Hospital de Torrecárdenas, Almería;

¹⁰Servicio de Medicina Interna, Hospital Vall d'Hebron, Barcelona;

¹¹Servicio de Medicina Interna, Hospital Virgen del Camino, Pamplona;

¹²Servicio de Medicina Interna, Hospital Universitari Mútua, Terrassa;

¹³Servicio de Medicina Interna, Hospital Universitario Carlos Haya, Málaga;

¹⁴Servicio de Reumatología, Hospital de la Princesa, Madrid;

¹⁵Instituto de Parasitología y Biomedicina López Neyra, Consejo Superior de Investigaciones Científicas, Granada.

References

1. MENDES D, CORREIA M, BARBEDO M *et al.*: Behçet's disease – a contemporary review. *J Autoimmun* 2009; 32: 178-88.
2. SAKANE T, TAKENO M, SUZUKI N, INABA G: Behçet's disease. *N Engl J Med* 1999; 341: 1284-91.
3. KONÉ-PAUT I, GEISLER I, WECHSLER B *et al.*: Familial aggregation in Behçet's disease: high frequency in siblings and parents of pediatric probands. *J Pediatr* 1999; 135: 89-93.
4. MENTHON DE M, LAVALLEY MP, MALDINI C, GUILLEVIN L, MAHR A: HLA-B51/B5 and the risk of Behçet's disease: a systematic review and meta-analysis of case-control genetic association studies. *Arthritis Rheum* 2009; 61: 1287-96.
5. YAZICI H, FRESKO I, YURDAKUL S: Behçet's syndrome: disease manifestations, management and advances in treatment. *Nat Clin Pract Rheumatol* 2007; 3: 148-55.
6. REMMERS EF, COSAN F, KIRINO Y *et al.*: Genome-wide association study identifies variants in the MHC class I, IL10, and IL23R-IL12RB2 regions associated with Behçet's disease. *Nat Genet* 2010; 42: 698-702.
7. MIZUKI N, MEGURO A, OTA M *et al.*: Genome-wide association studies identify IL23R-IL12RB2 and IL10 as Behçet's disease susceptibility loci. *Nat Genet* 2010; 42: 703-6.
8. HATEMI G, SEYAH E, FRESKO I, TALARICO R, HAMURYUDAN V: Behçet's syndrome: a critical digest of the 2013-2014 literature. *Clin Exp Rheumatol* 2014; 32 (Suppl. 84): S112-22.
9. KANG EH, CHOI JY, LEE YJ, LEE EY, LEE EB, SONG YW: Single nucleotide polymorphisms in IL-10-mediated signalling pathways in Korean patients with Behçet's disease. *Clin Exp Rheumatol* 2014; 32: S27-32.
10. TURUNE G, COSKUN D, ALIBAZ-ONER F *et al.*: TIR-domain-containing adaptor protein gene TIRAP S180L polymorphism is not increased in Behçet's disease patients in two ethnic cohorts. *Clin Exp Rheumatol* 2013; 31: 54-6.
11. LEE YJ, HORIE Y, WALLACE GR *et al.*: Genome-wide association study identifies GIMAP as a novel susceptibility locus for Behçet's disease. *Ann Rheum Dis* 2013; 72: 1510-6.
12. ORTIZ-FERNÁNDEZ L, CONDE-JALDÓN M, GARCÍA-LOZANO JR *et al.*: GIMAP and Behçet's disease: no association in the European population. *Ann Rheum Dis* 2014; 73: 1433-4.
13. BETTENCOURT A, LEAL B, CARVALHO C *et al.*: CC chemokine receptor polymorphism CCR5Δ32 in Portuguese Behçet's disease patients. *Clin Exp Rheumatol* 2014; 32 (Suppl. 84): S72-4.
14. WU Z, CHEN H, SUN F *et al.*: PTPN22 rs1893217 single-nucleotide polymorphism is associated with risk of Behçet's disease in a Chinese Han population. *Clin Exp Rheumatol* 2014; 32: S20-6.
15. BABON JJ, LUCET IS, MURPHY JM, NICOLA NA, VARGHESE LN: The molecular regulation of Janus Kinase (JAK) activation. *Biochem J* 2014; 462: 1-13.
16. BOONE DL, TURER EE, LEE EG *et al.*: The ubiquitin-modifying enzyme A20 is required for termination of Toll-like receptor responses. *Nat Immunol* 2004; 5: 1052-60.
17. DIEUDÉ P, GUEDJ M, WIPFF J *et al.*: Association of the *TNFAIP3* rs5029939 variant with systemic sclerosis in the European Caucasian population. *Ann Rheum Dis* 2010; 69: 1958-64.
18. ZHANG X, LI W, ZHANG X *et al.*: Single nucleotide polymorphisms in *TNFAIP3* were associated with the risks of rheumatoid arthritis in northern Chinese Han population. *BMC Medical Genetics* 2014; 15: 56.
19. HU K, HOU S, LI F, XIANG Q, KILSTRA A, YANG P: JAK1, but not JAK2 and STAT3, confers susceptibility to Vogt-Koyanagi-Harada (VKH) syndrome in a Han Chinese population. *Invest Ophthalmol Vis Sci* 2013; 54: 3360-5.
20. HOU S, QI J, ZHANG Q *et al.*: Genetic variants in the *JAK1* gene confer higher risk of Behçet's disease with ocular involvement in Han Chinese. *Hum Genet* 2013; 132: 1049-58.
21. LI H, LIU Q, HOU S *et al.*: *TNFAIP3* gene polymorphisms confer risk for Behçet's disease in a Chinese Han population. *Hum Genet* 2013; 132: 293-300.
22. INTERNATIONAL STUDY GROUP FOR BEHÇET'S DISEASE: Criteria for diagnosis of Behçet's disease. *Lancet* 1990; 335: 1078-80.
23. MONTES-CANO MA, CONDE-JALDÓN M, GARCÍA-LOZANO JR *et al.*: HLA and non-HLA genes in Behçet's disease: a multicentric study in the Spanish population. *Arthritis Res Ther* 2013; 15: R145.
24. PURCELL S, NEALE B, TODD-BROWN K *et al.*: PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; 81: 559-75.
25. KARASNEH J, GÜL A, OLLIER WE, SILMAN AJ, WORTHINGTON J: Whole-genome screening for susceptibility genes in multicase families with Behçet's disease. *Arthritis Rheum* 2005; 52: 1836-42.
26. FEI Y, WEBB R, COBB BL, DİRESKENELİ H, SARUHAN-DİRESKENELİ, SAWALHA AH: Identification of novel genetic susceptibility loci for Behçet's disease using a genome-wide association study. *Arthritis Res Ther* 2009; 11: R66.
27. KIRINO Y, BERTSIAS G, ISHIGATSUBO Y *et al.*: Genome-wide association analysis identifies new susceptibility loci for Behçet's disease and epistasis between HLA-B*51 and ERAP1. *Nat Genet* 2013; 45: 202-7.

Association of CCR5Δ32 and Behçet's disease: new data from a case-control study in the Spanish population and meta-analysis

L. Ortiz-Fernández¹, J.-R. García-Lozano¹, M.-A. Montes-Cano¹, M. Conde-Jaldón¹, N. Ortego-Centeno², M.-J. Castillo-Palma³, G. Espinosa⁴, G. Graña-Gil⁵, J. Sánchez-Bursón⁶, M.R. Juliá⁷, R. Blanco⁸, A.-C. Barnosi-Marín⁹, R. Solans¹⁰, P. Fanlo¹¹, M. Rodríguez Carballeira¹², T. Camps¹³, S. Castañeda¹⁴, J. Martín¹⁵, M.-F. González-Escribano¹

Lourdes Ortiz-Fernández,
José-Raúl García-Lozano,
Marco-Antonio Montes-Cano,
Marta Conde-Jaldón,
Norberto Ortego-Centeno,
María-Jesús Castillo-Palma,
Gerard Espinosa, Genaro Graña-Gil,
Juan Sánchez-Bursón, María Rosa Juliá,
Ricardo Blanco, Ana-Celia Barnosi-Marín,
Roser Solans, Patricia Fanlo,
Mónica Rodríguez Carballeira,
Teresa Camps, Santos Castañeda,
Javier Martín,

María-Francisca González-Escribano

Authors' affiliations on page 5.

This work should be attributed to:
Servicio de Inmunología, IBI-S, Hospital
Universitario Virgen del Rocío/CSIC/
Universidad de Sevilla, Sevilla. Spain.

Please address correspondence to:
María Francisca González-Escribano,
Servicio de Inmunología,
HU Virgen del Rocío,
41013 Sevilla, Spain.

E-mail:
mariafgonzalez.sspa@juntadeandalucia.es

Received on February 3, 2015; accepted in
revised form on June 4, 2015.

© Copyright CtinCa1 and
ExpErimEnta1 r hEumato1ogy 2015.

Key words: Behçet disease, CCR5
gene, CCR5 delta 32 deletion, genetic
contribution, meta-analysis.

Funding: this work was supported by
Fondo de Investigaciones Sanitarias
(10/1701 and 13/01118), Fondos FEDER,
Plan Andaluz de Investigación (CTS-0197),
Red Enfermedades Inflamatorias y
Reumáticas (RD08/0075/0013) and
Consejería de Salud de la Junta de
Andalucía (PI411/10).

L. Ortiz-Fernández is the recipient of a
fellowship (FI11/00547).

Competing interests: none declared.

ABSTRACT

Objective. Behçet's disease (BD) is an immune-mediated and complex disease associated with HLA class I and other genes. The aim of this study was to contribute to a better understanding of the relationship of the 32-bp deletion in the CCR5 gene (CCR5Δ32) and this disease by conducting a case-control study in the Spanish population and also a meta-analysis including all the studies available to date.

Methods. A cohort composed of 348 BD Spanish patients and 477 unrelated healthy and ethnically matched individuals were genotyped in CCR5Δ32 using polymerase chain reaction (PCR) and capillary electrophoresis with fluorescent detection. In the meta-analysis, data from a total of seven populations extracted from four previous studies along with data of the present study were included.

Results. Regarding the case-control study, no statistically significant differences were observed when the patient and control groups were compared (allelic model: 0.07 in patients vs. 0.06 in controls, $p=0.303$). In the meta-analysis, no evidence of association of the CCR5Δ32 polymorphism with BD was observed ($p_{MH}=0.091$; $OR=1.22$; 95%CI 0.98 to 1.52 in the allelic model).

Conclusion. The results of this meta-analysis discard a major role of the CCR5Δ32 polymorphism in BD.

Introduction

Behçet's disease (BD) is a multisystemic inflammatory disorder characterised by recurrent oral and genital ulcerations; ocular affection, mainly uveitis, and skin lesions, additional manifesta-

tions which involving other organs such as joints and central nervous system are relatively common. The aetiology of BD remains elusive, although it has been suggested that the disease is the result of complex interactions between environmental factors (e.g. certain infectious agents) and genetic predisposition (1).

The evidences of genetic contribution to the pathogenesis of the disease are based on familial aggregation, predominance in patients with Mediterranean or Asian ancestry and the association with human histocompatibility complex (HLA) in several ethnic groups (2-4). The contribution of the HLA region has been estimated to represent approximately 20% of the genetic component of this disease therefore other genes should be involved in the predisposition to this pathology (5). As a result of different studies designed to establish the contribution of genes located outside the HLA class I region with the disease, a relationship of BD with IL23R, IL10 and other genes has been established in different populations (6-12).

One of the non-HLA genes proposed as a candidate in BD is the chemokine receptor type 5 (CCR5), a G protein-coupled receptor which is expressed on Th1 cells, monocytes and dendritic cells. Binding of CCR5 to its ligands mediates the migration of mononuclear cells to the inflammation site and, in addition, the CCR5 molecule is also a co-receptor of the HIV. The most extensively studied polymorphism in this gene is a deletion of 32 bp (CCR5Δ32) which results in the introduction of a premature stop codon by a reading frameshift producing, as a consequence, a truncated protein unable to

bind to the natural ligands of CCR5 (13). The *CCR5Δ32* allele is related to the resistance to HIV infection (14) and it has also been associated with immune-based diseases such as rheumatoid arthritis and multiple sclerosis (15, 16). The role of the *CCR5Δ32* variant in immuno-mediated diseases is not clear because in some conditions such as rheumatoid arthritis it appears as a protective factor (17) because it is associated with a reduced risk whereas in others such as multiple sclerosis it increases the risk to develop the disease (18). CCR5 mediates mononuclear cell recruitment to sites of inflammation by interacting with its ligands CCL3, CCL4 and CCL5. The lack of functional CCR5 causes up-regulation of its ligands which can exert their biological effects by engaging other available receptors (19). Thus, lymphocytes from *CCR5Δ32* homozygous subjects secrete CCL5 at levels that are 5–10 times higher than the CCR5 non-mutated control subjects (20). The elevated levels of this chemokine can, by engaging other available receptors such as CCR3 and CCR1 result in increased recruitment of inflammatory cells and production of pro-inflammatory cytokines.

So far, four studies investigating the association between CCR5 gene and BD have been published (21–24). Three of these studies found no association of the *CCR5Δ32* variant with susceptibility to BD. Nevertheless, they include cohorts with a relatively small sample size and this fact, along with the low frequency of the variant, determines that the statistical power of the individual studies is inadequate. The aim of this study was to contribute to improve the current knowledge about the relationship of this functional variant of the CCR5 gene and BD by investigating whether the *CCR5Δ32* variant is associated with BD in the Spanish population and also by conducting a meta-analysis including all the available data.

Material and methods

Case-control study

This part of the study includes 348 BD patients (153 males and 195 females) with a mean age at onset (years) \pm SD of 48.22 ± 12.19 who fulfilled the 1990

International Study Group classification criteria for Behçet's disease (25) and 477 unrelated healthy individuals recruited in the same geographic regions and matched by age and gender with BD patients. All the subjects were Spanish from European origin, patient group was partially included in the Registry of the Spanish network of Behçet's disease (26) and they were recruited from different Spanish hospitals. The study was approved by the local ethics committees of all the participant hospitals., A Coruña (CHU A Coruña), Almería (H. Torredcardenas), Barcelona (H. Clinic, Vall d'Hebron and Mútua Terrassa), Granada (H. Clínico San Cecilio), Madrid (H. de la Princesa), Málaga (H. Carlos Haya), Palma de Mallorca (H. Universitari Son Espases), Pamplona (H. Virgen del Camino), Santander (H. Marques de Valdecilla) and Sevilla (H. Virgen del Rocío y H. Virgen de Valme) and a written informed consent was obtained from all participants. Clinical features of the patient group were: 100% had oral ulcers, 66% genital ulcers, 59% uveitis, 51.6% arthritis, and 22% vascular, 23% neurological and 20% gastrointestinal involvement. Peripheral blood or saliva were used as the starting material. Genomic DNA was extracted using the QIAamp DNA Mini Kit (Quiagen, Barcelona, Spain) according to the manufacturer's recommendations and stored at -20°C until use. The purity of DNA was determined using NanoDrop spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA). Only those DNA samples having a 260/280 ratio between 1.7 and 2.0 and a final concentration of 10–20 ng/μl were genotyped.

Genotyping was performed by polymerase chain reaction (PCR) using primers spanning the region of the 32 bp deletion in the *CCR5* gene. The sequence of the primers were those previously described 5'CTTCATCATCTCCTGACAAATCG3' (sense) and 5'GAC-CAGCCCAAGTTGACTATC3' (antisense) (27), labelling the sense primer with a fluorochrome (Fam) at the 5' end. The PCR products were separated by capillary electrophoresis using a 3130/3130xl Genetic Analyzer (Applied Biosystems, Barcelona, Spain)

with fluorescent detection. The expected size of the amplified DNA fragments for the wild-type and deletion were 262 bp and 230 bp respectively. About 5% of the samples were studied in duplicate to verify genotyping consistency which was identical in all the cases.

The statistical powers were calculated with the Statistical Power test of the Research Tool Kit of the DSS Research Software (<https://www.dssresearch.com/KnowledgeCenter/toolkitcalculators/statisticalpowercalculators.aspx>) to detect associations with odds ratios (OR) greater than 1.5 in the allelic model and taking into account the allelic frequency of the variant in each population. All the statistical analyses were performed with PLINK V.1.07 (<http://pngu.mgh.harvard.edu/purcell/plink/>). The χ^2 test was used to test Hardy Weinberg equilibrium (HWE). To check association of the variant *CCR5Δ32* with disease, the distribution of alleles and genotypes in patient and control groups were compared using χ^2 (or Fisher exact test when appropriated) from 2x2 contingency tables. Odds ratios (ORs) and 95% confidence intervals (CI 95%) were calculated according to the Woolf method.

Meta-analysis

We conducted a meta-analysis which includes the data of the present study together with those previously published regarding the relationship of the *CCR5Δ32* variant and BD. Literature included in the analysis was selected using the PubMed database (<http://www.ncbi.nlm.nih.gov/pubmed/>) and Scopus (<http://www.scopus.com>) searching for “CCR5 or chemokine receptor type 5” and “polymorphism” and “Behçet's disease”. References from the retrieved papers were also checked. Papers published up to January 2015 were considered to be included in the meta-analysis. The eligible studies were included when they met the following criteria: association studies written in English, in which commonly accepted classification criteria for Behçet's disease were used, reporting data of the population under study and displaying genotype distribution for the *CCR5Δ32* polymorphism both in the patient and control groups. Exclusion criteria were HWE

deviation, insufficient information and redundant or overlapping results.

The combined data were summarised in two by two tables. The Brewslo-Day method, the Cochran χ^2 -based Q test and the inconsistency index (I^2) were used to assess heterogeneity in the different populations. Brewslo-Day and Q test with p -values (P_{BD} and P_O) lower than 0.05 were considered statistically significant. Those I^2 values from 0 to 25% were considered non heterogeneity, 25–50% moderate, 50–75% large and 75–100% extreme heterogeneity. The ORs were pooled using the Mantel-Haenszel, Robins-Breslow-Greenland methods for fixed effects because non significant heterogeneity was detected and p -values (P_{MH}) lower than 0.05 were considered significant. All associations were tested under the allelic and the dominant model and the meta-analysis was conducted using StatsDirect v.2.6.6 (StatsDirect, Altrincham, UK) software. In addition, to examine the degree to which an individual study affects the overall estimate, sensitivity analyses were conducted by removing one study at a time and analysing the change of the pooled effect.

Results

Case-control study

In the cohort included in our study, the successful rate of genotyping was 95%, the study population was found to be in the Hardy-Weinberg equilibrium ($p>0.05$) and distribution of the *CCR5Δ32* variant in the cohorts from different hospitals was not significantly different. Table I displays the data of genotyping of *CCR5Δ32* in our cohort of BD patients and healthy controls. Frequency of the *CCR5Δ32* variant in our control population was 0.06 similar to that described in other South European populations and lower than that found in the North of Europe where the highest frequency of the *CCR5Δ32* variant has been described. No statistically significant differences were observed when the patient and control groups were compared under the allelic model (0.07 in patients vs. 0.06 in controls, $p=0.303$). Regarding the genotypes distribution, no significant differences were found comparing the distribution in pa-

Table I. Frequency of the *CCR5Δ32* genotypes in Spanish BD patients and healthy controls. The allelic and dominant models of inheritance are also displayed.

	Patients (n=348)		Controls (n=477)		p -value
Genotype	n	%	n	%	
Δ32/Δ32	1	0.3	0	0	0.325
Δ32/wt	46	13.2	54	11.3	
wt/wt	301	86.4	423	88.7	
Allelic Model					
Δ32	48	6.9	54	5.7	0.303
wt	648	93.1	900	94.3	
Dominant model					
Δ32/Δ32 + Δ32/wt	47	13.5	54	11.3	0.390
wt/wt	301	86.5	423	88.7	

wt: Wild type; Δ32: *CCR5Δ32* variant.

tients and controls under the dominant model ($p=0.390$). Statistical analyses of the other inheritance models were not performed because of the low number of individuals *CCR5Δ32/CCR5Δ32*. Regarding gender and HLA-B*51 subgroups, frequencies of the variant in our cohort were 0.07 in males vs. 0.06 in females ($p=0.688$) and 0.06 in the subjects B*51 positive vs. 0.07 in those negative ($p=0.551$). Therefore, no significant differences attributable to gender neither to the presence of B*51 were found. The possible association between the *CCR5Δ32* variant and the main clinical characteristics of BD were analysed but non significant differences were found (data not shown).

Meta-analysis

To avoid the low statistical power of the individual studies published to date regarding the relationship of the *CCR5Δ32* polymorphism and BD (see Table II), a meta-analysis was conducted. This meta-analysis includes the data pooled of the four papers previously published (all of them fulfilled the above mentioned criteria) together with data of the present study. A summary of the data of all these studies are displayed in Table II. No heterogeneity was detected in the analysis of this polymorphism in the homogeneity analysis ($P_{BD}=0.1881$, $P_Q=0.2089$ and $I^2=28.7\%$). Therefore, the subsequent meta-analysis was performed using the fixed effect method with the pooled data from a total of 1114 patients and 1786 healthy controls from seven different populations. The statistical power of

the pooled data to detect an association with $OR \geq 1.5$ is 93%. As a result of this meta-analysis, no evidence of association of the *CCR5Δ32* polymorphism with BD was observed ($P_{MH}=0.091$; $OR=1.22$; 95%CI 0.98 to 1.52 in the allelic model) (Fig. 1). The results obtained considering a dominant model were very similar to those obtained with the allelic (data not shown).

Discussion

The results of the present study suggest no association of the *CCR5Δ32* with the disease in the Spanish population. Most of the four individual studies published until the present regarding the association of this polymorphism have reported the same conclusion. In fact, only one study performed in an Italian cohort reported *CCR5Δ32* as a risk factor for BD (23). Although most of the studies discarded influence of this variant in the susceptibility to the disease, the statistical power of the individual studies, including the present, is low to reach a definitive conclusion. In fact, in all the populations except in the UK and the Portuguese populations (21–24) the frequency of the variant is higher in patients and therefore, differences could not reach significance as a consequence of limitation in size. Nevertheless, according to data of the meta-analysis presented in this study, which is well powered, a major role of the *CCR5Δ32* polymorphism in BD is discarded.

In some of the previous studies, authors suggested association of this variant with a specific subgroup. Thereby, in a study in an Iranian population,

Table II. Characteristics of the studies included in the meta-analysis.

Author	Year	Population	Genotyping methods	Sample Size		Allelic frequency of the $\Delta 32$ variant		% SP*	p-value	OR (95%CI)
				Cases	Controls	Cases	Controls			
Yang <i>et al.</i>	2003	UK	PCR and agarose gel electrophoresis	131	325	0.13	0.14	67.3	>0.05	
		Turkish		109	96	0.06	0.04	21.7	>0.05	
		Palestinian		110	98	0.02	0.005	9.0	>0.05	
Mojtahedi <i>et al.</i>	2006	Iranian	PCR and agarose gel electrophoresis	100	380	0.09	0.05	37.2	>0.05	2.28 (1.1-4.8)
Atzeni <i>et al.</i>	2012	Italian	PCR and agarose gel electrophoresis	196	180	0.06	0.03	25.0	0.02	
Bettencourt <i>et al.</i>	2015	Portuguese	PCR and agarose gel electrophoresis	122	230	0.07	0.08	27	>0.05	
Present study	2014	Spanish	PCR and capillary electrophoresis	348	477	0.07	0.06	68.2	>0.05	

*SP-Statistical power was calculated taking into account the allelic frequency and the sample size of each population and an expected OR ≥ 1.5 .

the authors suggested association between *CCR5* $\Delta 32$ and BD but only in the case of female patients (22). Nevertheless, data in our population, consistently with the study in the Portuguese population (24), do not confirm an influence of the *CCR5* $\Delta 32$ variant which would be dependent on gender because the distribution of the variant was not different in males and females. Additionally, data in the UK population, suggest that the frequency of the *CCR5* $\Delta 32$ variant is increased among patients B*51 positive (21). This fact could reflect association only among individuals B*51 and/or a higher homogeneity of the patients with this HLA risk factor. On this sense, a previous meta-analysis studying the influence of this *CCR5* variant in BD and other diseases, suggested association

of the *CCR5* $\Delta 32$ among B*51 carriers (28). Nevertheless, this association was not found in the Portuguese population (24) neither in the present study in the Spanish population. The global distribution of this *CCR5* variant is well known mainly because of its interest in the outcome of the HIV infection. In general, this mutation is more common among people with European ancestry but in the southern Europe, including Mediterranean populations, the frequency is usually lower than in the northern Europe. Frequencies that are found in the Middle East populations are similar to those found in the Mediterranean area and its presence is sporadic in the Asian populations (29). For this reason, although BD has a relatively high incidence in the Japanese and Chinese populations the influence

of this variant in susceptibility to this disease in Asian populations did not have investigated because this variant practically does not exist in these populations. The question about association of *CCR5* remains open in Asian populations because the influence of other functional variants of this gene which are present in these populations has not been yet investigated.

In conclusion, according to data of the meta-analysis presented in this study, a major role of the *CCR5* $\Delta 32$ polymorphism in BD can be discarded=

Acknowledgments

The authors would like to thank the Asociación Andaluza de Enfermedades Autoinmunes (AADEA) and all patients and donors enrolled in the present study for their cooperation.

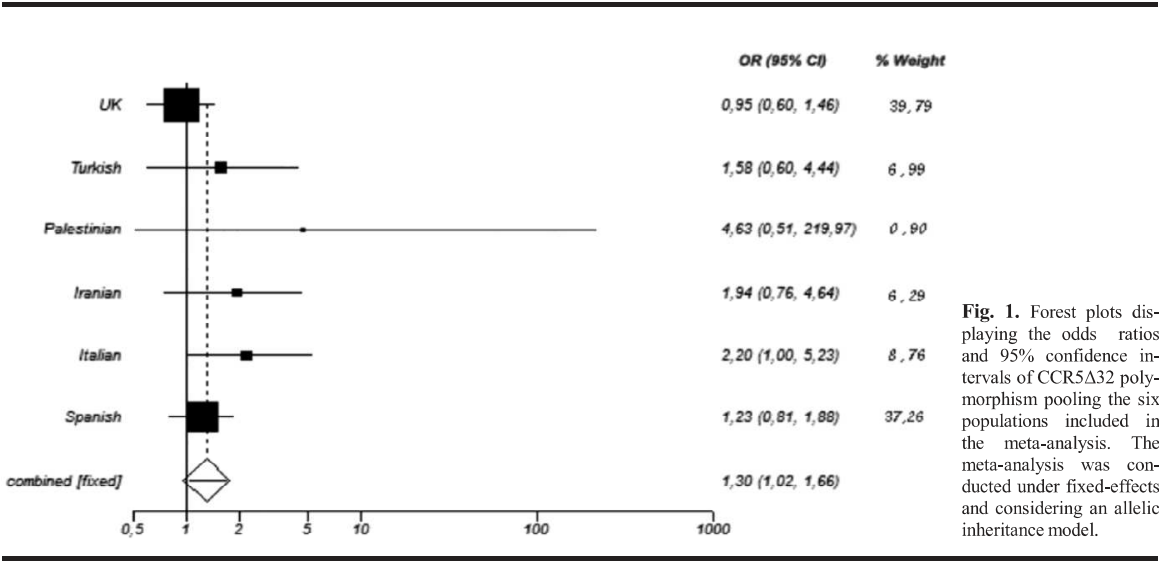


Fig. 1. Forest plots displaying the odds ratios and 95% confidence intervals of *CCR5* $\Delta 32$ polymorphism pooling the six populations included in the meta-analysis. The meta-analysis was conducted under fixed-effects and considering an allelic inheritance model.

Authors' affiliations

- ¹Servicio de Inmunología, IBiS, Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Sevilla;
- ²Servicio de Medicina Interna, Hospital Clínico San Cecilio, Granada;
- ³Servicio de Medicina Interna, Hospital Universitario Virgen del Rocío, Sevilla;
- ⁴Servicio de Enfermedades Autoinmunes, Hospital Clinic, Barcelona;
- ⁵Servicio de Reumatología, Complejo Hospitalario Universitario, A Coruña;
- ⁶Servicio de Reumatología, Hospital Universitario de Valme, Sevilla;
- ⁷Servicio de Inmunología, Hospital Universitari Son Espases, Palma de Mallorca;
- ⁸Servicio de Reumatología, Hospital Marqués de Valdecilla, Santander;
- ⁹Servicio de Medicina Interna, Hospital de Torrecárdenas, Almería;
- ¹⁰Servicio de Medicina Interna, Hospital Vall d'Hebron, Barcelona;
- ¹¹Servicio de Medicina Interna, Hospital Virgen del Camino, Pamplona;
- ¹²Servicio de Medicina Interna, Hospital Universitari Mútua, Terrassa;
- ¹³Servicio de Medicina Interna, Hospital Universitario Carlos Haya, Málaga;
- ¹⁴Servicio de Reumatología, Hospital de la Princesa, Madrid;
- ¹⁵Instituto de Parasitología y Biomedicina López-Neyra, Consejo Superior de Investigaciones Científicas, Granada, Spain.

References

- MENDES D, CORREIA M, BARBEDO M *et al.*: Behçet's disease: a contemporary review. *J Autoimmun* 2009; 32: 178-88.
- KONÉ-PAUT I, GEISLER I, WECHSLER B *et al.*: Familial aggregation in Behçet's disease: high frequency in siblings and parents of pediatric probands. *J Pediatr* 1999; 135: 89-93.
- SAKANE T, TAKENO M, SUZUKI N, INABA G: Behçet's disease. *N Engl J Med* 1999; 341: 1284-91.
- MENTHON de M, LAVALLEY MP, MALDINI C, GUILLEVIN L, MAHRA: HLA-B51/B5 and the risk of Behçet's disease: a systematic review and meta-analysis of case-control genetic association studies. *Arthritis Rheum* 2009; 61: 1287-96.
- YAZICI H, FRESKO I, YURDAKUL S: Behçet's syndrome: disease manifestations, management, and advances in treatment. *Nat Clin Pract Rheumatol* 2007; 3: 148-55.
- HUGHES T, COIT P, ADLER A *et al.*: Identification of multiple independent susceptibility loci in the HLA region in Behçet's disease. *Nat Genet* 2013; 45: 319-24.
- MEGURO A, INOKO H, OTA M *et al.*: Genetics of Behçet disease inside and outside the MHC. *Ann Rheum Dis* 2010; 69: 747-54.
- KARASNEH J, GÜL A, OLLIER WE, SILMAN AJ, WORTHINGTON J: Whole-genome screening for susceptibility genes in multi-case families with Behçet's disease. *Arthritis Rheum* 2005; 52: 1836-42.
- FEI Y, WEBB R, COBB BL, DİRESKENELI H, SARUHAN-DİRESKENELI, SAWALHA AH: Identification of novel genetic susceptibility loci for Behçet's disease using a genome-wide association study. *Arthritis Res Ther* 2009; 11: R66.
- REMMERS EF, COSAN F, KIRINO Y *et al.*: Genome-wide association study identifies variants in the MHC class I, IL10, and IL23R-IL12RB2 regions associated with Behçet's disease. *Nat Genet* 2010; 42: 698-702.
- MIZUKI N, MEGURO A, OTAM *et al.*: Genome-wide association studies identify IL23R-IL12RB2 and IL10 as Behçet's disease susceptibility loci. *Nat Genet* 2010; 42: 703-6.
- HATEMI G, SEYAHİ E, FRESKO I, TALARICO R, HAYAMURYUDAN V: Behçet's syndrome: a critical digest of the 2013-2013 literature. *Clin Exp Rheumatol* 2014; 32 (Suppl. 84): S112-122.
- CHARO IF, RANSOHOFF RM: The many roles of chemokines and chemokine receptors in inflammation. *N Engl J Med* 2006; 354: 610-21.
- LIU R, PAXTON WA, CHOE S *et al.*: Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* 1996; 86: 367.
- RODRÍGUEZ-RODRÍGUEZ L, GONZÁLEZ-JUANATEY C, GARCÍA-BERMÚDEZ M *et al.*: CCR5Δ32 variant and cardiovascular disease in patients with rheumatoid arthritis: a cohort study. *Arthritis Res Ther* 2011; 13: R133.
- PULKKINEN K, LUOMALA M, KUUSISTO H *et al.*: Increase in CCR5 Delta32/Delta32 genotype in multiple sclerosis. *Acta Neurol Scand* 2004; 109: 342-7.
- PRAHALAD S: Negative association between the chemokine receptor CCR5-Delta32 polymorphism and rheumatoid arthritis: a meta-analysis. *Genes Immun* 2006; 7: 264-8.
- FAVOROVA OO, ANDREEWSKI TV, BOIKO AN *et al.*: The chemokine receptor CCR5 deletion mutation is associated with MS in HLA-DR4-positive Russians. *Neurology* 2002; 26: 1652-5.
- CHARO IF, RANSOHOFF RM: The many roles of chemokines and chemokine receptors in inflammation. *N Engl J Med* 2006; 354: 610-21.
- PAXTON WA, KANG S: Chemokine receptor allelic polymorphisms: relationships to HIV resistance and disease progression. *Semin Immunol* 1998; 10: 187-94.
- YANG X, AHMAD T, GOGUS F *et al.*: Analysis of the CC chemokine receptor (CCR5) Delta32 polymorphism in Behçet's disease. *Eur J Immunogenet* 2004; 31: 11-4.
- MOJTAHEDI Z, AHMADI SB, RAZMKHAH M, AZAD TK, RAJAEI A, GHADERI A: Association of chemokine receptor 5 (CCR5) Δ32 mutation with Behçet's disease is dependent on gender in Iranian patients. *Clin Exp Rheumatol* 2006; 24: S91-4.
- ATZENI F, BOIARDI L, CASALI B *et al.*: CC chemokine receptor 5 polymorphism in Italian patients with Behçet's disease. *Rheumatology* 2012; 51: 2141-5.
- BETTENCOURT A, LEAL B, CARVALHO C *et al.*: CC chemokine receptor polymorphism CCR5Δ32 in Portuguese Behçet's disease patients. *Clin Exp Rheumatol* 2014; 32: S72-74.
- International Study group for Behçet's disease: Criteria for diagnosis of Behçet's disease. *Lancet* 1990; 335: 1078-80.
- RODRÍGUEZ-CARBALLEIRA M, ALBA MA, SOLANS-LAQUÉ R *et al.*: Registry of the Spanish network of Behçet's disease: a descriptive analysis of 496 patients. *Clin Exp Rheumatol* 2014; 32 (Suppl. 84): S33-9.
- KRISTIANSEN TB, KNUDSEN TB, OHLENDORFF S, EUGEN-OLSEN J: A new multiplex PCR strategy for the simultaneous determination of four genetic polymorphisms affecting HIV-1 disease progression. *J Immunol Methods* 2001; 252: 147-51.
- SONG GG, KIM JH, LEE YH: The chemokine receptor 5 delta32 polymorphism and type 1 diabetes, Behçet's disease, and asthma: a meta-analysis. *Immunol Invest* 2014; 43: 123-36.
- MARTINSON JJ, CHAPMAN NH, REES DC, LIU YT, CLEGG JB: Global distribution of the CCR5 gene 32-basepair deletion. *Nat Genet* 1997; 16: 100-3.

Fe de erratas

La figura 1 de la página 4 de este artículo es la que se muestra a continuación:

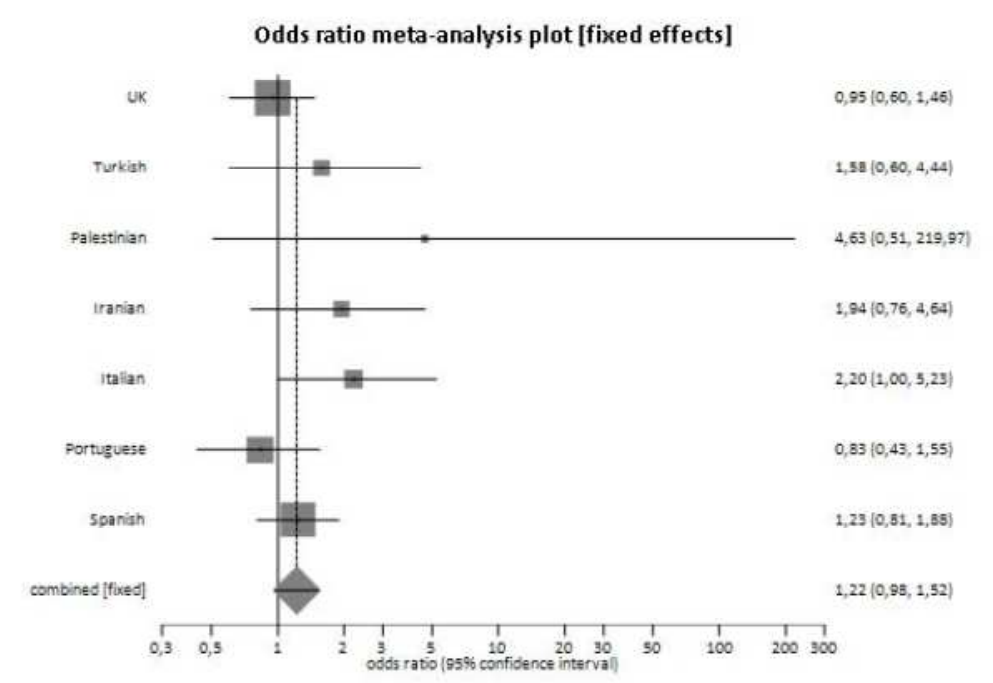


Fig 1. Forest plots displaying the odds ratios and 95% confidence intervals of CCR5 Δ 32 polymorphism pooling the seven populations included in the meta-analysis was conducted under fixed-effects and considering an allelic inheritance model.

PTPN22 is not associated with Behcet's disease. Study spanning the complete gene region in the Spanish population and meta-analysis of the functional variant R620W

Journal:	<i>The Journal of Rheumatology</i>
Manuscript ID	2015-1091
Manuscript Type:	Manuscript
Date Submitted by the Author:	15-Sep-2015
Complete List of Authors:	<p>Ortiz-Fernández, Lourdes; Hospital Universitario Virgen del Rocio, Servicio de Inmunología</p> <p>Montes-Cano, Marco-Antonio; Hospital Universitario Virgen del Rocio, Servicio de Inmunología</p> <p>García-Lozano, Jose-Raul; Hospital Universitario Virgen del Rocio, Servicio de Inmunología</p> <p>Conde-Jaldón, Marta; Hospital Universitario Virgen del Rocio, Servicio de Inmunología;</p> <p>Ortego, Norberto; Hospital Clínico San Cecilio, Internal Medicine Department</p> <p>Gonzalez-Leon, Rocio; Hospital Universitario Virgen del Rocio, Servicio de Medicina Interna</p> <p>Espinosa, Gerard; Hospital Clinic, Autoimmune Diseases</p> <p>Graña-Gil, Genaro; Complejo Hospitalario Universitario, Servicio de Reumatología</p> <p>Sanchez-Burson, Juan; HU Virgen de Valme, Servicio de Reumatología</p> <p>Julia, Maria; Son Espases University Hospital, Immunology Department</p> <p>Solans-Laque, Roser; Vall d'Hebron UNiversity Hospital, Internal Medicine</p> <p>Blanco, Ricardo; Hospital Universitario Marqués de Valdecilla, Rheumatology</p> <p>Barnosi-Marin, Ana Celia; HU Torrecardenas, Servicio de Medicina Interna</p> <p>Fanlo, Patricia; Hospital Virgen del Camino, Servicio de Medicina Interna</p> <p>Rodríguez-Carballeira, Mónica; Hospital Universitari Mútua Terrassa, Internal Medicine</p> <p>Camps, Teresa; Hospital Universitario Carlos Haya, Servicio de Medicina Interna</p> <p>Castañeda, Santos; , Hospital de la Princesa, Universidad Autónoma, Rheumatology</p> <p>Martin, Javier; Instituto de Parasitología y Biomedicina, Immunology</p> <p>González-Escribano, Maria; virgen del rocio university hospital, immunology</p>
Keywords:	Behcet's disease, Genetic predisposition to disease, Autoimmune diseases

For Peer Review

***PTPN22* is not associated with Behçet's disease. Study spanning the complete gene region in the Spanish population and meta-analysis of the functional variant R620W**

Lourdes Ortiz-Fernández¹, Marco-Antonio Montes-Cano¹, José-Raúl García-Lozano¹, Marta Conde-Jaldón¹, Norberto Ortego-Centeno², Rocío González-León³, Gerard Espinosa⁴, Genaro Graña-Gil⁵, Juan Sánchez-Bursón⁶, María Rosa Juliá⁷, Roser Solans⁸, Ricardo Blanco⁹, Ana-Celia Barnosi-Marín¹⁰, Patricia Fanlo¹¹, Mónica Rodríguez Carballeira¹², Teresa Camps¹³, Santos Castañeda¹⁴, Javier Martín¹⁵, María Francisca González-Escribano^{1*}.

Abstract

Objectives: The functional variant R620W of the protein tyrosine phosphatase non receptor-22 (*PTPN22*) gene plays an important role in susceptibility to several immun-mediated pathologies. Behçet's disease (BD) is a complex disease related to the immune system with a demonstrated genetic base. The HLA class I genes are the most important genetic factors in BD although other genes are also involved in the susceptibility to this disease. The *PTPN22* has been proposed as a candidate gene in BD but this association has not been clearly demonstrated yet. The aim of this study was asses the association of *PTPN22* with BD.

Methods: A cohort composed of 404 Spanish BD patients and 1517 unrelated healthy individuals ethnically matched was genotyped in rs2476601 (R620W). Five tag SNPs: rs1217412, rs2476599, rs3789607, rs3765598 and rs1217419 (spanning a 57 Kb region between 3'UTR and 5'UTR) and rs2488457 (located at the promoter region) were also studied in order to perform a screening of the complete gene. Genotyping was performed using TaqMan® assays. The rs2476601 data were included in a meta-analysis together with those published till the date. The rest of SNPs were used in a case-control study.

Results: No evidence of the association of rs2476601 with BD in the meta-analysis ($P = 0.504$ in the model of alleles) were found. In the case-control study, no statistically significant differences were observed when comparing the distribution of variants in patients and controls.

Conclusions: Our results discard a major role of the *PTPN22* gene in BD.

Key indexing terms: Behçet disease, *PTPN22*, R620W, genetic, meta-analysis

This work should be attributed to Servicio de Inmunología, IBiS, Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Sevilla.

Funding statement This work was supported by Fondo de Investigaciones Sanitarias (13/01118), Fondos FEDER and Plan Andaluz de Investigación (CTS-0197). LOF is the recipient of a fellowship (FI11/00547).

L Ortiz-Fernández PhD¹, Marco-Antonio Montes-Cano PhD¹, JR García-Lozano PhD,MD¹, M Conde-Jaldón PhD¹, N Ortego-Centeno MD², R González-León MD³, G Espinosa MD⁴, G Graña-Gil MD⁵, J Sánchez-Bursón MD⁶, M R Juliá PhD⁷, R Solans MD⁸, R Blanco MD⁹, AC Barnosi-Marín MD¹⁰, P Fanlo MD¹¹, M Rodríguez Carballeira MD¹², T Camps MD¹³, S Castañeda MD¹⁴, J Martín PhD, MD¹⁵, M F González-Escribano PhD^{1*}.

¹Servicio de Inmunología, IBiS, Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Sevilla; ²Servicio de Medicina Interna, Hospital Clínico San Cecilio, Granada; ³Servicio de Medicina Interna, Hospital Universitario Virgen del Rocío, Sevilla; ⁴Servicio de Enfermedades Autoinmunes, Hospital Clinic, Barcelona; ⁵Servicio de Reumatología, Complejo Hospitalario Universitario, A Coruña; ⁶Servicio de Reumatología, Hospital Universitario de Valme, Sevilla; ⁷Servicio de Inmunología. Hospital Universitari Son Espases. Palma de Mallorca; ⁸Servicio de Medicina Interna, Hospital Vall d'Hebron, Barcelona; ⁹Servicio de Reumatología, Hospital Marqués de Valdecilla, Santander; ¹⁰Servicio de Medicina Interna, Hospital de Torrecárdenas, Almería; ¹¹Servicio de Medicina Interna, Hospital Virgen del Camino, Pamplona; ¹²Servicio de Medicina Interna, Hospital Universitari Mútua, Terrassa; ¹³Servicio de Medicina Interna, Hospital Universitario Carlos Haya, Málaga; ¹⁴Servicio

de Reumatología, Hospital de la Princesa, Madrid; ¹⁵Instituto de Parasitología y Biomedicina López Neyra, Consejo Superior de Investigaciones Científicas, Granada.

Corresponding author and reprint request: María Francisca González-Escribano, Servicio de Inmunología. HU Virgen del Rocío. 41013 Sevilla. Spain. mariaf.gonzalez.sspa@juntadeandalucia.es. Tel +34955013228 Fax +34955013221

Short running footline: *PTPN22* and Behcet's disease

For Peer Review

Introduction

Behçet's disease (BD) is an immune-mediated and complex disease in which certain environmental factors such as infectious agents are the triggers of the disease in genetically predisposed individuals [1]. This rare condition characterized by recurrent oral, genital ulcerations and other manifestations such as ocular affectation, mainly uveitis, is most common along the old route of "Silk Road", stretching from China to the Mediterranean area [2]. This particular geographical distribution in conjunction with the familial aggregation and association with HLA class I molecules (specifically with HLA-B51) are evidences supporting a genetic base into the pathogenesis of the disease [3,4]. The contribution of the HLA region to the genetic component has been estimated in approximately 20% [5]. Some genes such as IL23R, IL10 and others have been related with BD in different populations [6,7], whereas in other cases the association seems to be limited to one ethnic group or specific population [8,9].

The protein tyrosine phosphatase nonreceptor-22 (*PTPN22*) gene, which is located on 1p13.2, encodes a lymphoid protein tyrosine phosphatase known as Lyp. *PTPN22* is a good candidate to be involved in susceptibility to autoimmune diseases, in fact, a functional variant, C1858T (R620W), has been related with multiple immuno-mediated pathologies such as rheumatoid arthritis [10], systemic lupus erythematosus [11] and systemic sclerosis [12]. This variant causes disruption of the binding between Lyp and Csk which suppresses T cell activation and, consequently, individuals with the variant have, at least theoretically, hyperresponsive T Two previous studies have investigated the association between this variant of *PTPN22* gene and BD but without conclusive results [13,14]. The main problem is the low frequency of R620W variant in different populations and its virtual absence in others, such as Asian populations

(<http://hapmap.ncbi.nlm.nih.gov/>). This circumstance makes more difficult to evaluate the role of this gene in a rare condition such as BD. To our knowledge, no study covering the entire region of the gene has been performed in order to evaluate the relationship of the complete *PTPN22* gene with BD. The aim of the present study was to contribute to clarify the potential role of *PTPN22* in BD by investigating a) the contribution of the R620W variant to the susceptibility to the disease by performing a meta-analysis including data from our cohort together with those reported till the date and b) whether variants located in other positions of this gene contribute to the susceptibility to BD by performing a case-control study with tag SNPs.

Material and Methods

Study subjects

This study included a total of 404 BD patients (43.7% males) who fulfilled the 1990 International Study Group classification criteria for Behçet's disease [15]. Moreover, 1517 healthy individuals (43% males) were included as control group. All the subjects were Spanish from European origin and they were recruited from different Spanish hospitals. The study was approved for the local ethical committees of the corresponding hospitals and all the study participants gave written informed consent according to the declaration of Helsinki. Clinical features of the patient group were: 100% had oral ulcers, 59.5% genital ulcers, 54% uveitis, 42% arthritis, and 21% vascular, 18.2% neurological and 15.5% gastrointestinal involvement. The distribution of the frequencies of the different markers in the cohorts from different hospitals was not significantly different.

DNA extraction

Peripheral blood or saliva were used as starting material. Genomic DNA was extracted using the QIAamp DNA Mini Kit (Qiagen, Barcelona, Spain) according to the manufacturer's recommendations and stored at -20°C until use. The purity of DNA was determined using a NanoDrop spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA). Only those DNA samples with a 260/280 ratio between 1.7 and 2.0 and a final concentration of 10-20 ng/µl were genotyped.

Genotyping of rs2476601 and meta-analysis

Genotyping of the rs2476601 was performed using TaqMan® SNP Genotyping Assays (Applied Biosystems, Barcelona, Spain) in a 7900HT Fast Real Time PCR System (Applied Biosystems). A meta-analysis, which included the results of genotyping in our cohort together with those previously published in studies that investigated the relationship of this variant of the *PTPN22* gene and BD was conducted.. References included in the analysis were selected from the databases PubMed (<http://www.ncbi.nlm.nih.gov/pubmed/>.) and Scopus (<http://www.scopus.com>) using the searching criteria: “PTPN22” or “protein tyrosine phosphatase non-receptor 22” and “polymorphism” and “Behçet's disease”. The date of the last revision was July 2015 and the references included in the retrieved papers were also checked. One study was considered as eligible when they met the following criteria: association studies written in English, in which commonly accepted classification criteria for Behçet's disease were used, containing information of the population under study and where the genotype distribution in the patient and control groups were displayed in detail. Exclusion criteria were deviation of Hardy-Weinberg equilibrium (HWE), insufficient information and redundant or overlapping results. One of eligible studies was excluded because of insufficient information [16]

Tag SNPs selection and genotyping

A *PTPN22* region (Chr1:114158106..114215631) LD plot was generated with the “Utah residents with Northern and Western European ancestry from the CEPH collection” (CEU) data obtained from HapMap (data release 28, Phase II+III, NCBI build 36 assembly, dbSNP b126) (<http://hapmap.ncbi.nlm.nih.gov/>). The confidence intervals approach was used to define haplotype blocks in the CEU population using the Haploview program (Haploview, V.4.1. Broad Institute, Cambridge, MA, USA). With these conditions, five tag SNPs (minor allele frequency MAF>0.1, r^2 threshold>0.8) were selected to genotype: rs1217412, rs2476599, rs3789607, rs3765598 and rs1217419 in addition to rs2476601. These six tag SNPs span a 57 Kb region between 3'UTR and 5'UTR in the *PTPN22* gene capturing the 27 SNPs in the region (100% coverage). In addition, rs2488457, which is located in the promoter region, was also included in the study because it is associated with different pathologies [17]. Genotyping was performed using TaqMan® SNP Genotyping Assays (Applied Biosystems) in a 7900HT Fast Real Time PCR System (Applied Biosystems).

Statistical Analysis

The software CaTS Power Calculator (<http://www.sph.umich.edu/csg/abecasis/CaTS/>) was used for calculating the statistical powers [18]. The χ^2 test was used to test HWE. The different association models: genotypic, allelic, dominant and recessive were tested using the χ^2 test and p-values \leq 0.05 were considered statistically significant. The ORs and 95% confidence intervals (95% CI) were calculated according to Woolf's method. For the meta-analysis, the combined data were summarized in two by two tables. The Brewslo-Day method, the Cochran χ^2 -based Q test and the inconsistency index (I^2) were used to assess heterogeneity in the different populations. Brewslo-Day and Q test

with P-values (P_{BD} and P_Q) lower than 0.05 were considered statistically significant. Those I^2 values from 0 to 25% were considered non heterogeneity, 25–50% moderate, 50–75% large and 75–100% extreme heterogeneity. The ORs were pooled using the DerSimonian-Laird method for random effects because heterogeneity was detected and P-values (P_{DL}) lower than 0.05 were considered significant. The associations were tested under the allelic model. In addition, to examine the extent to which an individual study affects the overall estimate, sensitivity analysis were performed to study how the elimination of one study each time affects to the pooled data. The Linux software PLINK V.1.07 (<http://pngu.mgh.harvard.edu/purcell/plink/>) and StatsDirect V.2.6.6 (StatsDirect, Altrincham, UK) software were used to carry out statistical analyses.

Results

The rs2476601 genotyping success rate was higher than 95% and no divergence from HWE was observed nor in controls neither in cases ($p > 0.05$). The allelic frequencies in the control group ($MAF = 0.064$) were similar to those reported in the 1000 Genomes Project for the Iberian population in Spain (IBS, 0.08) and Tuscan in Italy (TSI, 0.066) (<http://browser.1000genomes.org>). The meta-analysis, conducted with the data pooled of this study together with those included in two papers previously published, incorporates a total of 794 patients and 1876 controls from four different populations: United Kingdom (UK), Middle East (ME), Turkey and Spain. A summary of the data of all these studies are displayed in Table 1. To note that, two separate studies [13,14] detected significant differences in two different populations (UK and Turkey), whereas in other two independent studies, the present and other [13], performed also in two different populations (ME and Spain), significant differences were not observed. Heterogeneity across the different studies was detected in the analysis of this

polymorphism ($P_{BD}=0.002$, $P_Q=0.021$ and $I^2=69.1\%$), therefore, the meta-analysis was performed using the random effect method. No evidence of association of the rs2476601 (R620W) with BD was observed in the meta-analysis ($P_{DL}=0.504$; OR=0.723; 95%CI 0.28 to 1.87 in the allelic model) (Figure 1).

Next, Table 2 displays the allele and genotype frequencies in BD patients and healthy controls for the six tag SNPs (including the rs2476601) which were studied in Chr1:114158106..114215631 in order to check the whole gene region from 3'UTR until 5'UTR. This table includes also data of the rs2488457, which is located at the promoter region, and also the statistical power values for all the SNPs included considering three different ORs: 1.3, 1.5 and 2.0. The allelic frequencies in the control group were similar to those reported in the 1000 Genomes Project for IBS and TSI populations (<http://browser.1000genomes.org>). No statistically significant differences were observed when distributions in patients and controls were compared.

Discussion

The *PTPN22* gene, specifically its variant C1858T (R620W), is a very interesting candidate in the susceptibility to immune mediated diseases; in fact, it has been described in association with many of these pathologies [10-12]. Nevertheless, results of the present study discard a major role of this gene in BD. Although two independent studies performed in populations from UK and Turkey suggested association of the MAF of this variant of the *PTPN22* as a protective factor in BD, the p-values were only marginally significant [13,14]. Replication studies are essential to avoid type I errors, the present study includes the largest cohort of patients and controls genotyped for the variant R620W in BD and it is well powered to detect association with similar OR as the original study, nevertheless, no significant association was detected. Although

differences among populations can not be completely ruled out, the origin of the discrepancies may be also, the use of different typing methods because these have evolved from less reproducible methods, to others more reliable. Limitations caused by genotyping methods are more important in small cohorts, in which the Type I errors are more likely.

BD is a rare condition, this fact conditioned the sample size and it makes more difficult to achieve an adequate statistical power in individual studies (Table 1). This disadvantage is compounded because this variant, which has a low worldwide frequency, is extremely rare or absent in populations with a relatively high prevalence of the disease such as Turkish, Japanese and Chinese. For this reason a meta-analysis by pooling all the available data for this variant in BD was performed. The meta-analysis with data pooled of the four populations has an adequate statistical power ($\geq 86\%$) to detect associations with $OR \geq 1.4$ and it supports non-association of the BD with this variant of the *PTPN22* gene.

Other SNPs of this gene have been associated with immune mediated diseases in populations in which the C1858T (R620W) is absent. Thereby, rs2488457, rs1310182 and rs3789604 have been associated with type 1 diabetes and rheumatoid arthritis in Asian populations [19,20], nevertheless, association with BD has been discarded in the Chinese Han population in a large cohort of patients with ocular involvement [16]. Because of the interest of this gene in the immune mediated diseases and in order to investigate whether other variants are associated with BD, a study of tagSNPs was performed. This study, that spanning the whole *PTPN22* gene region with 100% coverage of all the 27 SNPs located within this zone and including in addition the rs2488457 (located at the promoter region), discards association of this gene with BD,

at least for $ORs \geq 1.3$, for which the study was adequately powered. In agreement with our results, none of the three GWAS performed in BD have reported association with *PTPN22* [6-7,21]. Although complete information derived from GWAS is not always available, in those cases where it has been reported, no association of this gene with BD was found [6]. In conclusion, our results of case-control study and meta-analysis discard a major role of the *PTPN22* in Behçet's disease.

Acknowledgments

The authors would like to thank to Asociación Andaluza de Enfermedades Autoinmunes (AADEA) and all patients and donors enrolled in the present study for their cooperation.

References

1. Mendes D, Correia M, Barbedo M, Vaio T, Mota M, Gonçalves O, Valente J. Behçet's disease: a contemporary review. *J Autoimmun* 2009; 32:178-88.
2. Sakane T, Takeno M, Suzuki N, Inaba G. Behçet's disease. *N Engl J Med* 1999; 341:1284-1291.
3. Koné-Paut I, Geisler I, Wechsler B, Ozen S, Ozdogan H, Rozenbaum M, Touitou I. Familial aggregation in Behçet's disease: high frequency in siblings and parents of pediatric probands. *J Pediatr* 1999; 135:89-93.
4. Menthon de M, Lavalley MP, Maldini C, Guillevin L, Mahr A. HLA-B51/B5 and the risk of Behçet's disease: a systematic review and meta-analysis of case-control genetic association studies. *Arthritis Rheum* 2009; 61:1287-96.
5. Yazici H, Fresko I, Yurdakul S. Behçet's syndrome: disease manifestations, management, and advances in treatment. *Nat Clin Pract Rheumatol* 2007; 3:148-55

6. Remmers EF, Cosan F, Kirino Y, Ombrello MJ, Abaci N, Satorius C, *et al.* Genome-wide association study identifies variants in the MHC class I, IL10, and IL23R-IL12RB2 regions associated with Behçet's disease. *Nat Genet* 2010; 42:698-702.
7. Mizuki N, Meguro A, Ota M, Ohno S, Shiota T, Kawagoe T, *et al.* Genome-wide association studies identify IL23R-IL12RB2 and IL10 as Behçet's disease susceptibility loci. *Nat Genet* 2010; 42:703-706.
8. Lee YJ, Horie Y, Wallace GR, Choi YS, Park JA, Choi JY, *al.* Genome-wide association study identifies GIMAP as a novel susceptibility locus for Behcet's disease. *Ann Rheum Dis* 2013; 72:1510-6.
9. Ortiz-Fernández L, Conde-Jaldón M, García-Lozano JR, Montes-Cano MA, Ortego-Centeno N, Castillo-Palma MJ, *et al.* GIMAP and Behçet disease: no association in the European population. *Ann Rheum Dis* 2014; 73:1433-4.
10. Totaro MC, Tolusso B, Napolioni V, Faustini F, Canestri S, Mannocci A, *et al.* PTPN22 1858 C>T Polymorphism distribution in Europe and association with rheumatoid arthritis: Case-Control Study and meta-Analysis. *PLoS ONE* 2011; 6:e24292.
11. Kyogoku C, Langefeld CD, Ortmann WA, Lee A, Selby S, Carlton VE, *et al.* Genetic association of the R620W polymorphism of protein tyrosine phosphatase PTPN22 with human SLE. *Am J Hum Genet* 2004; 75:504-7.
12. Dieudé P, Guedj M, Wipff J, Avouac J, Hachulla E, Diot E, *et al.* The PTPN22 620W allele confers susceptibility to systemic sclerosis: findings of a large case-control study of European Caucasians and a meta-analysis. *Arthritis Rheum* 2008; 58:2183-88.

13. Baranathan V, Stanford MR, Vaughan RW, Kondeatis E, Graham E, Fortune F, *et al.* The association of PTPN22 620W polymorphism with Behçet's disease. *Ann Rheum Dis* 2007; 66:1531-33.
14. Sahin N, Bicakcigil M, Atagunduz P, Direskeneli H, Saruhan-Direskeneli G. PTPN22 gene polymorphism in Behçet's disease. *Tissue Antigens* 2007; 70:432-34.
15. International Study group for Behçet's disease. Criteria for diagnosis of Behçet's disease. *Lancet* 1990; 335:1078-80.
16. Zhang Q, Hou S, Jiang Z, Du L, Li F, Xiao X, *et al.* No Association of PTPN22 Polymorphisms with Susceptibility to Ocular Behçet's Disease in Two Chinese Han populations. *PLoS ONE* 2012;7:e31230.
17. Chen Z, Zhang H, Xia B, Wang P, Jiang T, Song M, Wu J. Association of PTPN22 gene (rs2488457) polymorphism with ulcerative colitis and high levels of PTPN22 mRNA in ulcerative colitis. *Int J Colorectal Dis* 2013; 28:1351-8.
18. Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication based analysis for two-stage genome-wide association studies. *Nat Genet* 2006; 38:209-13.
19. Kawasaki E, Awata T, Ikegami H, Kobayashi T, Maruyama T, Nakanishi K, *et al.* Systematic search for single nucleotide polymorphisms in a lymphoid tyrosine phosphatase gene (PTPN22): association between a promoter polymorphism and type 1 diabetes in Asian populations. *Am J Med Genet A* 2006; 140:586-93.
20. Feng X, Li YZ, Zhang Y, Bao SM, Tong DW, Zhang SL, Hu CJ. Association of the PTPN22 gene (-1123G > C) polymorphism with rheumatoid arthritis in Chinese patients. *Tissue Antigens* 2010; 76:297-300.

21. Karasneh J, Gül A, Ollier WE, Silman AJ, Worthington J. Whole-genome screening for susceptibility genes in multicase families with Behçet's disease. *Arthritis Rheum* 2005; 52:1836-1842.

For Peer Review

Figure Legends

Figure 1. Forest plots of *PTPN22* rs2476601. The Forest plots of the rs2476601 in the *PTPN22* gene pooling the four populations included in the meta-analysis was performed with the DerSimonian-Laird method for random effects because heterogeneity was detected. The inheritance model considered was the allelic inheritance model. The odds ratios and 95% confidence intervals are displayed.

For Peer Review

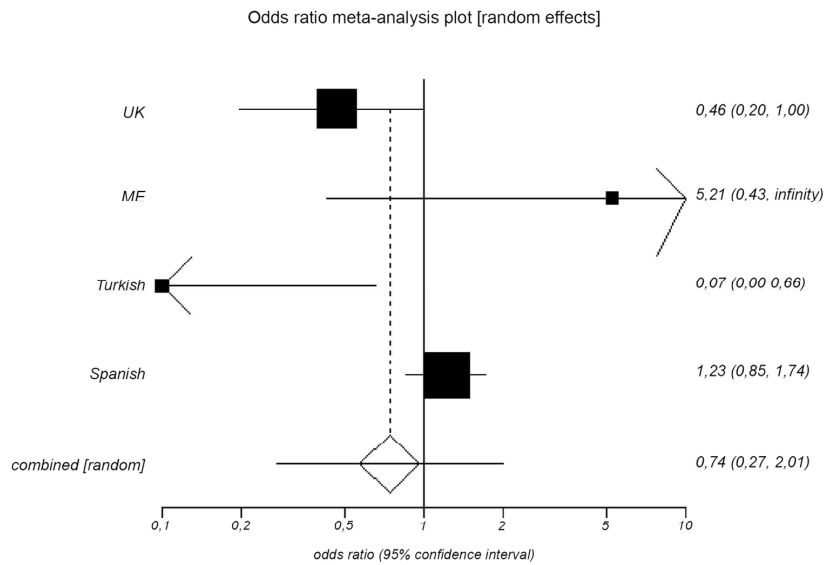


Figure 1
129x99mm (300 x 300 DPI)

Table 1. Characteristics of the studies included in the meta-analysis of the rs2476601

First author	Year	Population	Genotyping methods	Sample Size		MAF (T)		Statistical power ¹		P-value	OR
				Cases	Controls	Cases	Controls	OR=1.5	OR=2.0		
Baranathan V	2007	UK ²	PCR-SSP	116	169	4.5%	9.7%	30%	60%	0.04	0.46 ³
		ME		140	65	1.8%	0%	<10%	<10%	0.180	5.2
Sahin N	2007	Turkish	PCR + RFLP	134	177	0%	2.5%	15%	30%	0.012	0.65
Ortiz-Fernández L	2014	Spanish	Real Time PCR (TaqMan)	404	1465	6.7%	6.4%	76%	100%	0.723	1.06
Overall				794	1876					0.504	

¹Statistical powers were calculated for OR values of 1.5 and 2.0 taking into account the allelic frequency and the sample size of each population

²UK [-AC/ME]; UK cohort without patients from Afro-Caribbean or Middle Eastern (ME) origin

³The OR (95% IC) shown in the paper (2.1) corresponds with the most common allele. (C)

MAF= Minor Allele Frequency

Table 2. Distribution of genotypes and alleles of the SNPs included in this study

SNP	1/2	Subject	Genotype N (%)			MAF	Allele test		Statistical power		
			1/1	1/2	2/2		p-value*	OR (CI 95%)	OR 1.3	OR 1.5	OR 2.0
rs1217412	A/G	BD	162 (58.3)	105 (37.8)	9 (3.2)	0.223	0.863	0.98 (0.79-1.22)	81%	99%	100%
		Controls	903 (59.5)	517 (34.1)	81 (5.3)	0.226					
rs2476599	A/G	BD	16 (5.8)	115 (41.4)	145 (52.2)	0.266	0.571	1.06 (0.86-0.30)	84%	100%	100%
		Controls	99 (6.5)	565 (37.2)	833 (54.9)	0.255					
rs3789607	A/G	BD	123 (44.2)	122 (43.9)	33 (11.9)	0.338	0.763	1.03 (0.85-1.25)	88%	100%	100%
		Controls	676 (44.6)	676 (44.6)	165 (10.9)	0.332					
rs2476601	C/T	BD	346 (86.5)	54 (13.5)	0 (0)	0.067	0.723	1.06 (0.77-1.44)	38%	76%	100%
		Controls	1323 (87.2)	190 (12.5)	2 (0.1)	0.064					
rs3765598	A/G	BD	6 (2.2)	81 (29.1)	190 (68.3)	0.168	0.270	0.87 (0.69-1.11)	76%	99%	100%
		Controls	56 (3.7)	457 (30.1)	1003 (66.1)	0.188					

rs1217419	A/C	BD	37 (13.3)	145 (52.2)	96 (34.5)	0.394	0.406	0.92 (0.77-1.11)	100%	100%	100%
		Controls	266 (17.5)	717 (47.3)	530 (34.9)	0.413					
rs2488457	C/G	BD	185 (66.5)	88 (31.7)	5 (1.8)	0.176	0.447	0.91 (0.72-1.16)	76%	99%	100%
		Controls	998 (65.8)	460 (30.3)	58 (3.8)	0.190					

*All the p-values were calculated in the allelic model

MAF= Minor Allele Frequency

Variants of the *IFI16* Gene Affecting the Levels of Expression of mRNA Are Associated with Susceptibility to Behçet Disease

Lourdes Ortiz-Fernández, José-Raúl García-Lozano, Marco-Antonio Montes-Cano, Marta Conde-Jaldón, Norberto Ortego-Centeno, Francisco-José García-Hernández, Gerard Espinosa, Genaro Graña-Gil, Juan Sánchez-Bursón, Ricardo Blanco, Ana-Celia Barnosi-Marín, Roser Solans, Patricia Fanlo, Mónica Rodríguez-Carballeira, Teresa Camps, Santos Castañeda, Antonio Núñez-Roldán, Javier Martín, and María-Francisca González-Escribano

ABSTRACT. Objective. Behçet disease (BD) is a multifactorial disease in which infectious agents have been proposed as triggers in genetically predisposed individuals. The aim of our study was to investigate the role of innate immunity receptors, specifically the nucleic acid sensors, in susceptibility to BD. **Methods.** Seventy-four tag single nucleotide polymorphisms (tSNP) selected in 9 candidate genes (*DDX58*, *IFIH1*, *TLR3*, *TLR7*, *TLR8*, *AIM2*, *IFI16*, *ZBP1*, and *TLR9*) were genotyped in 371 patients and 854 controls. Assays of mRNA expression and allele-specific transcript quantification (ASTQ) were performed in 110 and 50 controls, respectively. **Results.** Patients and controls were genotyped and 2 tSNP (rs6940 in *IFI16* and rs855873 in *AIM2*) were associated with BD. To confirm this association, these tSNP were genotyped in 850 additional controls, and the total cohort was randomly divided into 2 cohorts. The association of these 2 tSNP with the disease remained in both cohorts. One haplotype (rs6940T-rs855873G) was identified as a risk factor (OR 1.41, 95% CI 1.06–1.86, $p = 0.015$), and another (rs6940A-rs855873A) as a protective factor (OR 0.65, 95% CI 0.47–0.90, $p = 0.009$). Samples with the risk haplotype had lower *IFI16* expression levels than samples with the protective (0.99 ± 0.29 vs 1.23 ± 0.50 , $p = 0.022$). Consistently, in the ASTQ assays performed with the nonsynonymous rs6940 SNP, the risk allele had lower *IFI16* expression levels than the protective ($p = 0.027$). **Conclusion.** Our findings suggest association of *IFI16*, a cytosolic sensor of dsDNA and mediator of the AIM2 inflammasome-dependent pathway, in susceptibility to BD. Differences genetically determined in the levels of this molecule could be the cause of this association. (J Rheumatol First Release Feb 1 2015; doi:10.3899/jrheum.140949)

Key Indexing Terms:
BEHÇET DISEASE *IFI16* *AIM2* INFLAMMASOME

From the Servicio de Inmunología, IBI^S, Hospital Universitario Virgen del Rocío/Consejo Superior de Investigaciones Científicas (CSIC)/Universidad de Sevilla, Sevilla; Servicio de Medicina Interna, Hospital Clínico San Cecilio, Granada; Servicio de Medicina Interna, Hospital Universitario Virgen del Rocío, Sevilla; Servicio de Enfermedades Autoinmunes, Hospital Clinic, Barcelona; Servicio de Reumatología, Complejo Hospitalario Universitario, La Coruña; Servicio de Reumatología, Hospital Universitario de Valme, Sevilla; Servicio de Reumatología, Hospital Marqués de Valdecilla, Santander; Servicio de Medicina Interna, Hospital de Torrecárdenas, Almería; Servicio de Medicina Interna, Hospital Vall d'Hebron, Barcelona; Servicio de Medicina Interna, Hospital Virgen del Camino, Pamplona; Servicio de Medicina Interna, Hospital Universitari Mútua, Terrassa; Servicio de Medicina Interna, Hospital Universitario Carlos Haya, Málaga; Servicio de Reumatología, Hospital de la Princesa, Madrid; Instituto de Parasitología y Biomedicina López Neyra, CSIC, Granada, Spain.
Supported by Fondo de Investigaciones Sanitarias (10/1701), Fondos FEDER, Plan Andaluz de Investigación (CTS-0197 and CTS-180), Red Enfermedades Inflamatorias y Reumáticas RD08/0075/0013 and Consejería de Salud de la Junta de Andalucía (PI0411/2010). LOF is the recipient of a fellowship (FI11/00547).

L. Ortiz-Fernández, BSc; J.R. García-Lozano, MD, PhD; M.A. Montes-Cano, PhD; M. Conde-Jaldón, BSc; A. Núñez-Roldán, MD, PhD; M.F. González-Escribano, PhD; Servicio de Inmunología, Hospital Universitario Virgen del Rocío; N. Ortego-Centeno, MD, Servicio de Medicina Interna, Hospital Clínico San Cecilio; F.J. García-Hernández, MD, Servicio de Medicina Interna, Hospital Universitario Virgen del Rocío; G. Espinosa, MD, Servicio de Enfermedades Autoinmunes, Hospital Clinic; G. Graña-Gil, MD, Servicio de Reumatología, Complejo Hospitalario Universitario; J. Sánchez-Bursón, MD, Servicio de Reumatología, Hospital Universitario de Valme; R. Blanco, MD, Servicio de Reumatología, Hospital Marqués de Valdecilla; A-C. Barnosi-Marín, MD, Servicio de Medicina Interna, Hospital de Torrecárdenas; R. Solans, MD, Servicio de Medicina Interna, Hospital Vall d'Hebron; P. Fanlo, MD, Servicio de Medicina Interna, Hospital Virgen del Camino; M. Rodríguez-Carballeira, MD, Servicio de Medicina Interna, Hospital Universitari Mútua; T. Camps, MD, Servicio de Medicina Interna, Hospital Universitario Carlos Haya; S. Castañeda, MD, Servicio de Reumatología, Hospital de la Princesa; J. Martín, PhD, Instituto de Parasitología y Biomedicina López Neyra, CSIC.
Address correspondence to Dr. J.R. García-Lozano, Servicio de Inmunología, Hospital Universitario Virgen del Rocío, Avda Manuel

Behçet disease (BD) is a systemic vasculitis mainly characterized by recurrent oral and genital ulceration, although other clinical manifestations such as skin lesions, ocular, gastrointestinal, and neurological disorders are relatively common¹. The prevalence depends on the geographic location of the study population. The highest corresponds to Turkey (80 to 420 cases per 100,000 people), followed by Japan (70–85/100,000). Among Western countries, the Mediterranean basin has the highest prevalence (e.g., Spain has 0.7/100,000)^{1,2}. Although the etiology of BD remains unclear, it is known that it is an immune-mediated disease in which imbalances in the innate or adaptive immune response, triggered by infectious agents or environmental factors in genetically predisposed individuals, may be some of the underlying mechanisms of this multifactorial disease³. Regarding the environmental factors, infectious agents of the oral microbial flora (oral aphthae are the first manifestation in most patients) have been proposed as triggers of the disease⁴. Regarding the genetic factors, *HLA-B51* has been the most consistently associated with the disease. Nevertheless, the contribution of the *HLA* region represents about 20% of the genetic component^{5,6} and other non-*HLA* genes such as interleukin 23R (*IL-23R*) and *IL-10*, related to innate immunity, have been associated more recently with this pathology^{7,8}.

The innate immune system, which is the first barrier against pathogens, uses sensor molecules known as pattern recognition receptors (PRR) to recognize highly conserved structures expressed by microorganisms (pathogen-associated molecular patterns; PAMP) or by damaged cells (damage-associated molecular patterns; DAMP). Some of the PRR recognize components such as nucleic acids and therefore their activation creates, at least potentially, a risk of developing autoimmune disease⁹. Among the PRR involved in the recognition of nucleic acid molecules transported to the endosomal compartment by endocytosis or autophagy, Toll-like receptors (TLR) 3, TLR7, TLR8, and TLR9 are the best known. TLR3 recognizes double-stranded RNA (dsRNA), TLR7, and TLR8 single-stranded RNA (ssRNA) and TLR9 unmethylated dinucleotide CpG motifs of viral and bacterial DNA. After binding of the TLR to their ligands, the transduction signal produces activation of different molecules such as interferon regulatory factors (IRF) 3, IRF7, activating protein 1 and nuclear factor- κ B (NF- κ B), and at the end, transcription of inflammatory cytokines and type I interferon (IFN)¹⁰. Other sensors such as the RIG-I-like receptor (RLR) family, which include DDX58 and IFIH1, detect dsRNA in the cytosol. Similar to TLR, RLR induce activation of IRF3, IRF7, and NF- κ B and trigger the transcription of type I IFN¹¹. Besides the RNA sensors, a large number of sensors of intracellular DNA are

known. ZBP1 acts as a cytosolic dsDNA sensor that initiates IFN responses through activation of the NF- κ B and IRF3 pathways¹¹. IFI16 and AIM2 are cytosolic sensors for dsDNA members of AIM2-like receptors (ALR)^{12,13}; nevertheless, IFI16 and AIM2 proteins have different effects¹⁴. The IFI16 protein stimulates the expression of type I IFN (IFN- β) by activation of IRF3 and NF- κ B¹², and the AIM2 protein activates caspase-1 and leads to inflammation by the cleavage of inflammatory cytokine proforms, such as IL-1 β , IL-18, and IL-33^{15,16,17}. A schematic overview of the different aforementioned PRR, their ligands, and the signal transduction pathways is shown in Figure 1.

It has been proposed that certain infectious agents, especially those in the oral microbial flora, act as triggers of BD. Therefore, it is possible that variations in PRR influence the outcome of BD. The aim of our present study was to analyze the possible relationship between nucleic acid sensors in susceptibility to BD.

MATERIALS AND METHODS

Patients and controls. A total of 371 BD-unrelated patients (45% men) who fulfilled the 1990 International Study Group classification criteria for BD¹⁸ and 2 cohorts of 854 and 850 ethnically matched healthy unrelated bone marrow donors (50% men) were included in the study. All the subjects were Spanish whites recruited from different Spanish hospitals. All local ethics committees of the corresponding hospitals approved the study and all participants in the study gave written informed consent. Mean age at diagnosis was 38.7 ± 13.8 years and 100% of the patients had oral ulcers, 64% genital ulcers, 59% uveitis, and 48% arthritis. Twenty percent had vascular, 22% neurological, and 19% gastrointestinal involvement.

Peripheral blood was obtained from the healthy controls, whereas peripheral blood or saliva was the starting material obtained from patients. Genomic DNA was extracted using QIAamp DNA Mini Kit (Qiagen) according to the manufacturer's recommendations and stored at -20°C . The expression assays were performed using samples from bone marrow and blood donors with normal hematological parameters (hematocrit, platelet, lymphocytes, monocytes, and granulocytes). The cDNA was synthesized from total RNA purified from 107 peripheral blood mononuclear cells isolated by density gradient centrifugation as described¹⁹.

Single nucleotide polymorphism (SNP) selection and genotyping. Nine candidate genes were selected on the basis of their function in innate host defense as dsRNA (*DDX58*, *IFIH1*, *TLR3*), ssRNA (*TLR7*, *TLR8*), dsDNA (*AIM2*, *IFI16*, *ZBP1*), and CpG DNA (*TLR9*) sensors. Tag SNP (tSNP) were selected across each locus from the designated set of common SNP genotyped in the CEU population (HapMap Project, Release 28, Phase II + III, US National Center for Biotechnology Information build 36 assembly, dbSNP b126; www.hapmap.org). The tSNP selection was done with pairwise $r^2 \geq 0.80$ and minor allele frequency ≥ 0.05 using the Haploview v4.0 software (www.broad.mit.edu/mpg/haploview/download.php)²⁰.

According to the above rules, 74 tSNP that permitted us to identify 216 SNP were selected (Supplementary Table 1, available online at jrheum.org). All 371 patients with BD and 854 controls were genotyped using the SEQUENOM iPLEX MassARRAY platform according to manufacturer's instructions (Sequenom). Primers for multiplex PCR and extended reactions were designed using the MassARRAY Assay Design Suite v1.0 software (www.sequenom.com). Primers design was not possible in 8 tSNP that were genotyped using TaqMan SNP Genotyping Assays (Applied Biosystems) in a LightCycler 480 (Roche; Supplementary Table 1). To verify interexperimental reproducibility and accuracy, 8% of the samples were duplicated. A 90% sample quality control rate and 90% SNP genotyping success rate was imposed on the analysis. Additionally, the

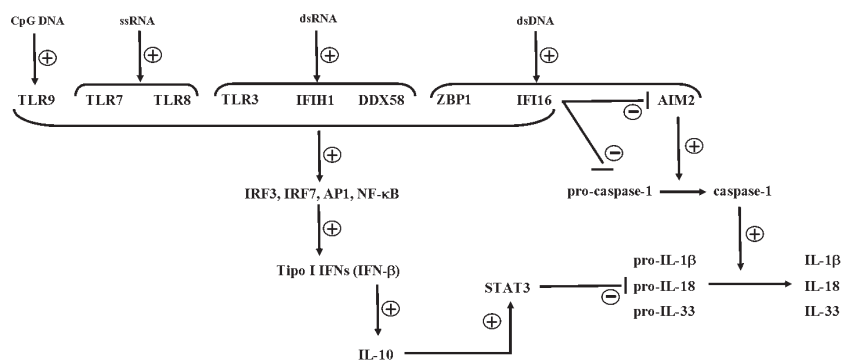


Figure 1. Pattern recognition receptors involved in the nucleic acid binding included in this study. TLR: Toll-like receptors; DDX58: DEAD (Asp-Glu-Ala-Asp) box polypeptide 58 (also known as retinoic acid inducible gene 1); IFIH1: interferon (IFN) induced with helicase C domain 1 (or melanoma differentiation-associated factor 5); ZBP1: Z-DNA binding protein 1 (or DNA-dependent activator of IFN regulatory factor); IFI16: IFN- γ -inducible protein 16; AIM2: absent in melanoma 2; IRF3, IRF7: IFN-regulatory factor 3 and 7; AP-1: activator protein 1; NF- κ B: nuclear factor- κ B; IL: interleukin; STAT1: signal transducers and activators of transcription-1.

Table 1. Minor allelic frequencies of the SNP associated with Behçet disease in Spanish patients.

Gene SNP	MA	Controls, n = 854	Patients, n = 371	p	p _c	OR (95% CI)
IFI16						
rs7532207	T	0.087	0.118	0.017	0.047	1.41 (1.06–1.85)
rs6940	T	0.088	0.120	0.013	0.040	1.42 (1.07–1.87)
AIM2						
rs855873	A	0.105	0.072	0.010	0.031	0.66 (0.48–0.91)

P values > 0.05 were not significant. SNP: single-nucleotide polymorphism; MA: minor allele; p_c: 10,000-fold permutation testing.

rs6940 and rs855873 SNP were genotyped in a second group of 850 controls using TaqMan SNP Genotyping Assays.

Quantification of mRNA AIM2 and IFI16. The AIM2 and IFI16 mRNA expression assays were performed using 110 samples from healthy controls on a LightCycler 480 by TaqMan Gene Expression Assay (Applied Biosystems). The mRNA expression of AIM2 and IFI16 was correlated with the mRNA expression of the reference ABL gene²¹. The conditions of the assays and data analysis have been described¹⁹. Each sample was tested in triplicate, and only samples showing ABL Cp values between 24 and 30 and an SD of Cp < 0.3 were included in the final analysis (n = 102).

Allele-specific transcript quantification. IFI16 mRNA allele-specific transcript quantification (ASTQ) assays were performed by TaqMan SNP Genotyping Assays using 50 samples from healthy controls who were heterozygotes for the nonsynonymous SNP rs6940. The Cp values for both alleles labeled 1 with FAM (rs6940A) and the other with VIC (rs6940T) and were obtained for cDNA (target) and gDNA (reference), and a relative quantification model was performed²¹. To detect the presence of contaminating gDNA in the RNA samples, aliquots of total RNA were previously amplified. Those RNA samples that did not pass this test were not processed. The samples processed were tested in triplicate and only those samples showing gDNA Cp values between 24 and 30 with a Cp SD < 0.3 were included in the final analysis (n = 46).

Statistical analysis. Allele and haplotype frequency distributions between patients with BD and controls were compared with the chi-squared test. The OR with their corresponding 95% CI were calculated using OpenEpi v2.3 software online (www.openepi.com). Statistical significance was evaluated using a 10,000-fold permutation test (p_c) and those p_c values <

0.05 were considered statistically significant. The genetic model of inheritance was evaluated with Chaplin v1.2 software (www.genetics.emory.edu/labs/epstein/software/chaplin), using the likelihood ratio statistics for testing the global null hypothesis²² and the Akaike information criterion to determine the genetic model of inheritance²³. Results of relative IFI16 and AIM2 mRNA expression were compared using an unpaired t test with Welch's correction when the variances were not homogeneous (Levene's test < 0.05). To compare values obtained by ASTQ, a paired sample t test was used. Statistical analyses were performed using SPSS v18.0 software (SPSS Inc.), and p values < 0.05 were considered statistically significant.

RESULTS

Three hundred seventy-one patients and 854 controls were genotyped in tSNP of 9 genes involved in the recognition of nucleic acid molecules: 13 tSNP in IFI16, 2 in AIM2, 4 in IFI1, 2 in TLR9, 8 in TLR3, 16 in DDX58, 9 in ZBP1, 12 in TLR7, and 8 in TLR8, capturing 56, 2, 4, 3, 14, 72, 13, 29, and 23 SNP, respectively (Supplementary Table 1, available online at jrheum.org). The results of 4 of the 74 tSNP included were discarded because they had amplification failures or were not meeting quality criteria. In the remaining 70 tSNP, the concordance of the assigned genotypes in duplicated samples was > 98% and the study population was found in Hardy-Weinberg equilibrium (p > 0.05). Two tSNP located in the IFI16 locus (rs7532207 and

rs6940) and 1 tSNP located in the *AIM2* locus (rs855873) were significantly associated with BD (Table 1). The minor allele of the rs855873 (A) was protective (7.2 vs 10.5% in the control group, $p = 0.010$, OR 0.66, 95% CI 0.48–0.91), whereas both the minor alleles of rs6940 (T) and rs7532207 (T) were risk alleles (12.0 vs 8.8% in the control group, $p = 0.013$, OR 1.42, 95% CI 1.07–1.87; and 11.8 vs 8.7% in the control group, $p = 0.017$, OR 1.41, 95% CI 1.06–1.85, respectively). The rs7532207 was not included in subsequent assays because a strong linkage disequilibrium ($D' = 1$ and $r^2 = 0.988$) between rs7532207 and rs6940 was found in our population in both patient and control groups.

To confirm the association between the *IFI16/AIM2* locus and BD, the rs6940 and rs855873 were genotyped in 850 additional healthy controls. Samples from patients and controls of the first cohort that failed to amplify using SEQUENON were genotyped with TaqMan assays as the new control cohort. The total cohort was randomly divided into 2 groups: the cohort 1, which consisted of 187 patients and 854 controls, and cohort 2, which included 184 patients and 850 controls. After re-analyzing, both cohorts retained significant association with susceptibility to BD ($p = 0.046$ and 0.042 to rs6940, and $p = 0.039$ and 0.027 to rs855873; Table 2).

IFI16 and *AIM2* genes are located in a region spanning 70-kb on chromosome 1q25.2 (Figure 2) and they are transcribed in opposite directions. By performing the analysis of the haplotypes constructed by combination of the rs6940 and rs855873, 3 combinations with frequency > 0.005 were identified in our population (Table 3). One haplotype, “TG,” which was tagged by the rs6940T allele, was identified as a risk haplotype, whereas another haplotype, “AA,” tagged by the rs855873A, was identified as protective. Testing different genetic models of inheritance, we found that the recessive model could be excluded for both the risk and the protective haplotypes, whereas the multiplicative ($p = 0.016$ and $p = 0.0033$, respectively) and the dominant ($p = 0.019$ and $p = 0.0058$, respectively) models were the best fitting (Supplementary Table 2, available online at jrheum.org).

The mRNA expression levels of the *AIM2* and the *IFI16* in 3 kinds of samples obtained from healthy individuals were compared: those having the risk haplotype (1 or 2 copies), those with the protective haplotype (1 or 2 copies), and finally, those homozygous for the neutral haplotype (individuals AG/AG; Table 4). Statistically significant differences in *AIM2* mRNA expression levels were not found; nevertheless, samples with the risk haplotype had lower *IFI16* mRNA expression than samples with the protective haplotype (0.99 ± 0.29 vs 1.23 ± 0.50 , $p = 0.022$), whereas samples homozygous for the neutral haplotype had an intermediate level of expression (1.08 ± 0.46). To confirm that the risk haplotype has the lowest *IFI16* mRNA expression levels, the nonsynonymous SNP rs6940 (T723S) was used to perform an ASTQ assay. Heterozygous samples from healthy individuals showed statistically significant lower *IFI16* mRNA expression for the risk (T) than for the protective (A) allele ($p = 0.027$) with a ratio A/T < 1 (0.97 ± 0.08 ; Supplementary Figure 1, available online at jrheum.org).

DISCUSSION

In our study, including a total of 74 tSNP in 9 candidate genes encoding nucleic acid sensors, we identified the *IFI16-AIM2* gene cluster as a novel genetic susceptibility locus for BD in a Spanish population. In addition, our data suggest that the steady-state basal *IFI16* mRNA expression could be haplotype-dependent.

BD is included among disorders related to inflammasomes because serum levels of IL-1 β , IL-18, and IL-33 are elevated among patients^{24,25,26} who respond to treatment with neutralizing antibodies for IL-1 β ²⁷. Several inflammasome-dependent pathways have been involved in BD pathogenesis. Thus, recently, production of IL-1 β through an NLRP3-inflammasome-dependent pathway, which is initiated with interaction of TLR2/TLR4 with their ligands, has been implicated in the pathogenesis of the disease²⁸. Our results suggest that the *AIM2* inflammasome could also contribute to the pathogenesis of BD, mainly characterized by the presence of recurrent oral and genital ulceration. It

Table 2. Replication study in 2 tSNP in *IFI16*: rs6940 and rs855873. The total cohort, composed of 371 patients and 1704 controls, was randomly divided into 2 groups: cohort 1 (187 patients and 854 controls) and cohort 2 (184 patients and 850 controls).

SNP	Cohort 1						Cohort 2						Cohort 1 + 2					
	Controls, n = 854		Patients, n = 187		p	OR (95% CI)	Controls, n = 850		Patients, n = 184		p	OR (95% CI)	Controls, n = 1704		Patients, n = 371		p	OR (95% CI)
	AF	NA	AF	NA			AF	NA	AF	NA			AF	NA	AF	NA		
rs6940-T	0.089	153	0.123	46	0.046	1.42 (1.01–2.02)	0.088	150	0.122	45	0.042	1.44 (1.02–2.05)	0.088	303	0.122	91	0.0045	1.43 (1.12–1.83)
rs-6940-A	0.911	1555	0.877	328			0.912	1550	0.878	323			0.912	3105	0.878	651		
rs855873-A	0.107	184	0.072	27	0.039	0.64 (0.42–0.98)	0.106	180	0.068	25	0.027	0.61 (0.40–0.95)	0.107	364	0.070	52	0.0025	0.63 (0.46–0.85)
rs855873-G	0.893	1524	0.928	347			0.894	1520	0.932	343			0.893	3044	0.930	690		

tSNP: tag single nucleotide polymorphisms; AF: allelic frequency; NA: no. alleles.

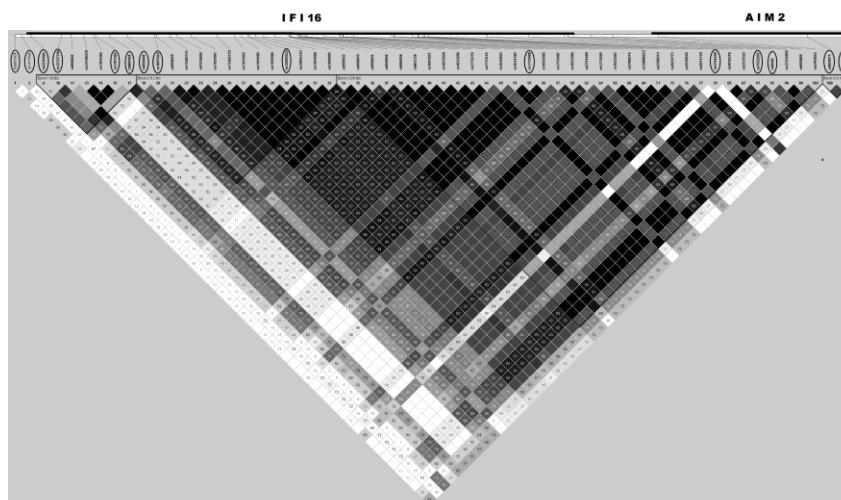


Figure 2. Haplotype blocks described in the region Chr1:157245220..157314336 of HapMap Project in the CEU population. The linkage disequilibrium (LD) solid spine approach was used to define the haplotype blocks. Standard color-coding was used for LD plots: white ($r^2 = 0$), shaded grey ($0 < r^2 < 1$), and black ($r^2 = 1$). Squares without a number indicate $D' = 1$. IFI16 and AIM2 loci are shown; marked rs numbers are the ones selected for this study.

Table 3. Allelic frequencies of the haplotypes constructed by combination of the 2 SNP associated with Behçet disease in Spanish patients.

	Haplotypes		Controls, n = 854	Patients, n = 371	P	P _c	OR (95% CI)
	rs6940	rs855873					
N	A	G	0.809	0.810	ns		
R	T	G	0.086	0.117	0.015	0.045	1.41 (1.06–1.86)
P	A	A	0.103	0.070	0.009	0.026	0.65 (0.47–0.90)

SNP: single nucleotide polymorphisms; N: neutral haplotype; R: risk haplotype; P: protective haplotype; ns: not significant ($p > 0.05$); p_c: 10,000-fold permutation testing.

Table 4. AIM2 and IFI16 mRNA expression levels in samples from healthy controls with different phenotypes.

Phenotypes	N	AIM2, Mean ± SD	p	IFI16, Mean ± SD	p
Risk	31	1.04 ± 0.57		0.99 ± 0.29	
Protective	35	0.99 ± 0.51	ns ¹	1.23 ± 0.50	0.022 ¹
Neutral	36	0.87 ± 0.47	ns ²	1.08 ± 0.46	ns ²

¹Risk versus protective. ²Risk versus neutral. Risk phenotypes are individuals bearing 2 risk haplotypes (TG/TG) or 1 risk and 1 neutral (TG/AG). Protective phenotypes are individuals bearing 2 protective haplotypes (AA/AA) or 1 protective and 1 neutral haplotype (AA/AG). Neutral phenotypes are individuals homozygous for the neutral haplotype (AG/AG). ns: not significant.

has been described as a strong upregulation of AIM2 protein expression in the epidermis of patients with several acute and chronic inflammatory skin disorders such as psoriasis, atopic dermatitis, venous ulcers, contact dermatitis, and in studies on wound healing^{29,30}. Possibly, AIM2 upregulation is a first line of defense against invading pathogens after skin barrier disruption. Nevertheless, this beneficial effect during wound healing could establish a chronic inflam-

matory injury in the skin when barrier impairment is prolonged because of activation induced by the IL-1 β AIM2³¹.

IFI16 and AIM2 are cytosolic sensors for dsDNA and both are members of the IFN-inducible p200 family (IFI200), which is also referred to as ALR³². IFI16 and AIM2 proteins have different effects because of their different pyrin domains (PYD), which allow protein-protein

interaction (Figure 1). The IFI16 protein, after recruiting STING (stimulator of interferon genes) through its PYD domain, stimulates the expression of IFN- β by activation of IRF3 and NF- κ B¹². Finally, IFN- β inhibits the production of IL-1 β by inflammasomes through a mechanism that is dependent on signal transducers and activators of transcription-1³³. For its part, the AIM2 protein, through its PYD domain, interacts with apoptosis-associated speck-like protein containing a CARD domain, which activates caspase-1 through its CARD domain¹³ and leads to inflammation by the cleavage of inflammatory cytokine proforms, such as IL-1 β , IL-18, and IL-33^{15,16,17}. Therefore, IFI16 has antiinflammatory effects, whereas AIM2 has proinflammatory effects. Moreover, IFI16 downregulates the expression of AIM2 and pro-caspase-1 and as a consequence, it suppresses the activation of caspase-1 by the AIM2 and NLRP3 inflammasomes¹⁴. Further, AIM2 and IFI16 can form heterodimers in the cytoplasm, and by this mechanism, IFI16 also downregulates the expression of an AIM2 inflammasome.

The IFI16/AIM2 ratio in the cytoplasm contributes to the balance between the production of type I IFN and IL-1 β ¹⁴. A study involving young, old, and senescent human diploid fibroblasts supports the theory that expression of IFI16 and AIM2 proteins is inversely correlated³⁴. Our data suggested that the steady-state basal *IFI16* mRNA expression could be haplotype-dependent because the risk haplotype has the lowest *IFI16* mRNA levels. These results were confirmed by performing ASTQ assays, which are not affected by confounding factors because mRNA levels for each allele are measured in the same sample of RNA and therefore differences in the mRNA expression levels for both alleles are a consequence of differences within the cis-acting regulatory elements. The analysis of expression was not performed in patients, to avoid problems related to different treatments (biologics and monoclonal therapy), which could modify the steady-state basal expression of the molecules under study.

Genome-wide association studies (GWAS) have identified several loci of susceptibility to BD outside the *HLA* region in different populations^{7,8,35,36,37,38}; nevertheless, association with the *IFI16* gene has not been described in any other population. Failure to identify this association in GWAS could be due to various reasons. The high-throughput platforms used in early GWAS for BD did not include the tSNP that we found associated in our study or other tSNP with $r^2 = 1$ with the rs6940 and rs855873^{7,8,35}. Although these tSNP are included in the high-throughput platforms of Affimatrix used in more recent studies^{36,37}, the sample size of the cohorts included was not large enough to detect the association with the threshold p-value levels used in GWAS.

Our study has limitations. The sample size allows only slightly significant results, and the low prevalence of the

disease in Spain does not permit recruiting many more patients. Therefore, these data need to be replicated in other populations with a higher prevalence of BD. However, it is also necessary to note that association of some non-*HLA* loci such as *IL10* and *IL23R-IL12RB* with the disease has been described in several populations, whereas other loci (such as *KIAA1529*, *CPVL*, *LOC100129342*, *UBASH3B*, *UBAC2*, *ERAP1*, *STAT4*, and *GIMAP*) have been found only in a particular ethnic group^{7,8,37,38,39}.

Our results suggest a relationship between the IFI16, a mediator of the AIM2 inflammasome-dependent pathway, and susceptibility to BD. This association could be explained by steady-state basal *IFI16* mRNA expression.

ACKNOWLEDGMENT

The authors thank Asociación Andaluza de Enfermedades Autoinmunes and all patients and donors enrolled in the present study for their cooperation. Also, the authors thank the following public Spanish institutions for their participation: Servicio Andaluz de Salud (Hospitales Torrecárdenas, Almería; Clínico San Cecilio, Granada; Carlos Haya, Málaga; and Virgen del Rocío and Valme, Sevilla); Servei Català de la Salut (Clinic and Vall d'Hebron hospitals, Barcelona); Servizo Galego de Saude (Complejo Hospitalario Universitario, A Coruña); Servicio Cántabro de Salud (Hospital Marqués de Valdecilla, Santander); Servicio Navarro de Salud (Hospital Virgen del Camino, Pamplona); Servicio Madrileño de Salud (Hospital de la Princesa, Madrid); and Consejo Superior de Investigaciones Científicas (IPB López Neyra, Granada).

ONLINE SUPPLEMENT

Supplementary data for this article are available online at jrheum.org.

REFERENCES

- Mendes D, Correia M, Barbedo M, Vaio T, Mota M, Goncalves O, et al. Behçet's disease—a contemporary review. *J Autoimmun* 2009;32:178-88.
- Gonzalez-Gay MA, Garcia-Porrúa C, Branas F, Lopez-Lazaro L, Olivieri I. Epidemiologic and clinical aspects of Behçet's disease in a defined area of Northwestern Spain, 1988-1997. *J Rheumatol* 2000;27:703-7.
- McGonagle D, McDermott MF. A proposed classification of the immunological diseases. *PLoS Med* 2006;3:e297.
- Lehner T. The role of heat shock protein, microbial and autoimmune agents in the aetiology of Behçet's disease. *Int Rev Immunol* 1997;14:21-32.
- de Menthon M, Lavalley MP, Maldini C, Guillemin L, Mahr A. HLA-B51/B5 and the risk of Behçet's disease: a systematic review and meta-analysis of case-control genetic association studies. *Arthritis Rheum* 2009;61:1287-96.
- Yazici H, Fresko I, Yurdakul S. Behçet's syndrome: disease manifestations, management, and advances in treatment. *Nat Clin Pract Rheumatol* 2007;3:148-55.
- Mizuki N, Meguro A, Ota M, Ohno S, Shiota T, Kawagoe T, et al. Genome-wide association studies identify IL23R-IL12RB2 and IL10 as Behçet's disease susceptibility loci. *Nat Genet* 2010;42:703-6.
- Remmers EF, Cosan F, Kirino Y, Ombrello MJ, Abaci N, Satorius C, et al. Genome-wide association study identifies variants in the MHC class I, IL10, and IL23R-IL12RB2 regions associated with Behçet's disease. *Nat Genet* 2010;42:698-702.
- Akira S, Uematsu S, Takeuchi O. Pathogen recognition and innate immunity. *Cell* 2006;124:783-801.

10. Kawasaki T, Kawai T, Akira S. Recognition of nucleic acids by pattern-recognition receptors and its relevance in autoimmunity. *Immunol Rev* 2011;243:61-73.
11. Takaoka A, Wang Z, Choi MK, Yanai H, Negishi H, Ban T, et al. DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response. *Nature* 2007;448:501-5.
12. Unterholzner L, Keating SE, Baran M, Horan KA, Jensen SB, Sharma S, et al. IFI16 is an innate immune sensor for intracellular DNA. *Nat Immunol* 2010;11:997-1004.
13. Hornung V, Ablasser A, Charrel-Dennis M, Bauernfeind F, Horvath G, Caffrey DR, et al. AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. *Nature* 2009;458:514-8.
14. Veeranki S, Duan X, Panchanathan R, Liu H, Choubey D. IFI16 protein mediates the anti-inflammatory actions of the type-I interferons through suppression of activation of caspase-1 by inflammasomes. *PLoS One* 2011;6:e27040.
15. Fernandes-Alnemri T, Yu J-W, Datta P, Wu J, Alnemri ES. AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA. *Nature* 2009;458:509-13.
16. Martinon F, Burns K, Tschopp J. The inflammasome: a molecular platform triggering activation of inflammatory caspases and processing of proIL-beta. *Mol Cell* 2002;10:417-26.
17. Arend WP, Palmer G, Gabay C. IL-1, IL-18, and IL-33 families of cytokines. *Immunol Rev* 2008;223:20-38.
18. International Study Group for Behçet's Disease. Criteria for diagnosis of Behçet's disease. *Lancet* 1990;335:1078-80.
19. Garcia-Lozano JR, Torres B, Fernandez O, Orozco G, Alvarez-Marquez A, Garcia A, et al. Caspase 7 influences susceptibility to rheumatoid arthritis. *Rheumatology* 2007;46:1243-7.
20. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263-5.
21. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 2001;29:e45.
22. Epstein MP, Satten GA. Inference on haplotype effects in case-control studies using unphased genotype data. *Am J Hum Genet* 2003;73:1316-29.
23. Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr* 1974;19:716-23.
24. Düzgün N, Ayaşlıoğlu E, Tutkac H, Aydıntuğ OT. Cytokine inhibitors: soluble tumor necrosis factor receptor 1 and interleukin-1 receptor antagonist in Behçet's disease. *Rheumatol Int* 2005;25:1-5.
25. Musabak U, Pay S, Erdem H, Simsek I, Pekel A, Dinc A, et al. Serum interleukin-18 levels in patients with Behçet's disease. Is its expression associated with disease activity or clinical presentations? *Rheumatol Int* 2006;26:545-50.
26. Hamzaoui K, Kaabachi W, Fazaa B, Zakraoui L, Mili-Boussen I, Haj-Sassi F. Serum IL-33 levels and skin mRNA expression in Behçet's disease. *Clin Exp Rheumatol* 2013;31:6-14.
27. Gül A, Tugal-Tutkun I, Dinarello CA, Reznikov L, Esen BA, Mirza A, et al. Interleukin-1 β -regulating antibody XOMA 052 (gevokizumab) in the treatment of acute exacerbations of resistant uveitis of Behçet's disease: an open-label pilot study. *Ann Rheum Dis* 2012;71:563-6.
28. Liang L, Tan X, Zhou Q, Zhu Y, Tian Y, Yu H, et al. IL-1 β triggered by peptidoglycan and lipopolysaccharide through TLR2/4 and ROS-NLRP3 inflammasome-dependent pathways is involved in ocular Behçet's disease. *Invest Ophthalmol Vis Sci* 2013;54:402-14.
29. Dombrowski Y, Peric M, Koglin S, Kammerbauer C, Göss C, Anz D, et al. Cytosolic DNA triggers inflammasome activation in keratinocytes in psoriatic lesions. *Sci Transl Med* 2011;3:82ra38.
30. de Koning HD, Bergboer JG, van den Bogaard EH, van Vlijmen-Willems IM, Rodijk-Olthuis D, Simon A, et al. Strong induction of AIM2 expression in human epidermis in acute and chronic inflammatory skin conditions. *Exp Dermatol* 2012;21:961-4.
31. Türe-Özdemir F, Tulunay A, Elbasi MO, Tatli I, Maurer A-M, Mumcu G, et al. Pro-inflammatory cytokine and caspase-1 responses to pattern recognition receptor activation of neutrophils and dendritic cells in Behçet's disease. *Rheumatology* 2013;52:800-5.
32. Veeranki S, Choubey D. Interferon-inducible p200-family protein IFI16, an innate immune sensor for cytosolic and nuclear double-stranded DNA: regulation of subcellular localization. *Mol Immunol* 2012;49:567-71.
33. Guarda G, Braun M, Staehli F, Tardivel A, Mattmann C, Förster I, et al. Type I interferon inhibits interleukin-1 production and inflammasome activation. *Immunity* 2011;34:213-23.
34. Duan X, Ponomareva L, Veeranki S, Panchanathan R, Dickerson E, Choubey D. Differential roles for the interferon-inducible IFI16 and AIM2 innate immune sensors for cytosolic DNA in cellular senescence of human fibroblasts. *Mol Cancer Res* 2011;9:589-602.
35. Fei Y, Webb R, Cobb BL, Direskeneli H, Saruhan-Direskeneli G, Sawalha AH. Identification of novel genetic susceptibility loci for Behçet's disease using a genome-wide association study. *Arthritis Res Ther* 2009;11:R66.
36. Hou S, Yang Z, Du L, Jiang Z, Shu Q, Chen Y, et al. Identification of a susceptibility locus in STAT4 for Behçet's disease in Han Chinese in a genome-wide association study. *Arthritis Rheum* 2012;64:4104-13.
37. Lee YJ, Horie Y, Wallace GR, Choi YS, Park JA, Choi JY, et al. Genome-wide association study identifies GIMAP as a novel susceptibility locus for Behçet's disease. *Ann Rheum Dis* 2013;72:1510-6.
38. Kirino Y, Bertsias G, Ishigatsubo Y, Mizuki N, Tugal-Tutkun I, Seyahi E, et al. Genome-wide association analysis identifies new susceptibility loci for Behçet's disease and epistasis between HLA-B*51 and ERAP1. *Nat Genet* 2013;45:202-7.
39. Ortiz-Fernandez L, Conde-Jaldon M, Garcia-Lozano JR, Montes-Cano MA, Ortego-Centeno N, Castillo-Palma MJ, et al. GIMAP and Behçet disease: no association in the European population. *Ann Rheum Dis* 2014;73:1433-4.

Table S1. Tag single-nucleotide polymorphisms (SNP) included in this study.

Gen	CHR	Tag SNP	SNP Identified	Genotyping
IFI16	1	rs9887904	rs9887904	SEQUENOM
		rs1417806	rs1417806	SEQUENOM
		rs2276404	rs2276404	SEQUENOM
		rs12727764	rs12727764	SEQUENOM
		rs1361564	rs12738378, rs1361565, rs1361564	SEQUENOM
		rs866484	rs866484, rs856064	SEQUENOM
		rs856046	rs856049, rs856046, rs11265131, rs7535856	SEQUENOM
		rs1057024	rs1057024	SEQUENOM
		rs12093523	rs12094741, rs1101991, rs12093523, rs11265133, rs856054, rs2106095, rs1101995, rs1101996, rs1101998, rs1101993	SEQUENOM
		rs1772408	rs855865, rs12124059, rs2570916, rs1772408, rs1772405, rs1057028, rs861318, rs856055, rs856052, rs856053, rs1633265, rs1772414, rs856058, rs1057027, rs1772415, rs1614182, rs1633256	SEQUENOM
		rs3754459	rs3754459	TaqMan
		rs7532207	rs7532207	SEQUENOM
		rs6940	rs6940, rs1772407, rs3018316, rs3737522, rs2814770, rs3768513, rs2814771, rs1616024, rs1633266, rs12098223, rs1633267, rs3754460, rs3768515	SEQUENOM
AIM2	1	rs855871	rs855871	SEQUENOM
		rs855873	rs855873	SEQUENOM
IFIH1	2	rs1990760	rs1990760	TaqMan
		rs3747517*	rs3747517	SEQUENOM
		rs13023380	rs13023380	SEQUENOM
		rs17715343	rs17715343	SEQUENOM
TLR9		rs352140	rs352140, rs352139	SEQUENOM
		rs187084	rs187084	SEQUENOM
TLR3	4	rs5743305	rs5743305	SEQUENOM
		rs11721827	rs11721827, rs11730143, rs11732384	SEQUENOM
		rs13126816	rs13126816, rs6552950	SEQUENOM
		rs13108688	rs13108688, rs7657186	SEQUENOM
		rs5743312	rs3775296, rs5743312, rs5743303	SEQUENOM
		rs7668666	rs7668666	SEQUENOM
		rs3775292	rs3775292	SEQUENOM
		rs3775291	rs3775291	SEQUENOM

Table S1, part 2.

DDX58	9	rs12006123	rs17289116, rs1133071, rs7045087, rs12555727, rs12006123	SEQUENOM
		rs7865082	rs7865082, rs3205166	TaqMan
		rs10738890	rs10738890, rs10738889	TaqMan
		rs9650702	rs6476362, rs9650702	SEQUENOM
		rs9695310	rs9695310	TaqMan
		rs944581	rs944581, rs639949, rs6476363, rs868608	SEQUENOM
		rs944582	rs944582	SEQUENOM
		rs17289402	rs17289402	SEQUENOM
		rs17289627	rs17289627, rs3824456, rs17290242	SEQUENOM
		rs944583	rs7037171, rs944583, rs4384073	SEQUENOM
		rs592515	rs10813821, rs7855688, rs10511907, rs10971001, rs592515, rs10813825, rs7848849	SEQUENOM
		rs13286888	rs13286888, rs13290848, rs13300627, rs13300238, rs11549548, rs7034650, rs17289655, rs12554841, rs13288658, rs17217280, rs17289927, rs669260	SEQUENOM
		rs10813826	rs10813826, rs582509, rs10970997	SEQUENOM
		rs608227	rs10738891, rs7026407, rs659527, rs668559, rs4633144, rs621277, rs1622531, rs613645, rs10813829, rs12340133, rs680471, rs3739674, rs10813827, rs608227, rs660819, rs2777729, rs7023269, rs7044989, rs605383, rs658253, rs7022323	SEQUENOM
		rs11795343	rs11795343	SEQUENOM
		rs10813831*	rs4013911, rs10970990, rs4013910, rs10813831	SEQUENOM
ZBP1	20	rs4811885*	rs4811885	SEQUENOM
		rs6064572	rs6064572, rs6123711	TaqMan
		rs8118279	rs8118279	TaqMan
		rs4811887	rs4811887, rs6070180, rs6025653	SEQUENOM
		rs4811888*	rs4811888	SEQUENOM
		rs742724	rs742724	SEQUENOM
		rs2073145	rs2073145	SEQUENOM
		rs16981188	rs16981188	SEQUENOM
		rs4811890	rs2865395, rs4811890	SEQUENOM
TLR7	X	rs5741880	rs5743733, rs5741880, rs5935436	TaqMan
		rs179021	rs179018, rs179021	SEQUENOM
		rs179019	rs179019, rs179014	SEQUENOM
		rs1731479	rs1634322, rs1731479	SEQUENOM
		rs179016	rs179016	SEQUENOM
		rs5743749	rs5743749	SEQUENOM
		rs1638596	rs1638596, rs1634321, rs1634320, rs1634323, rs1731478, rs1620233, rs1638597, rs1638595, rs1638594	SEQUENOM
		rs179013	rs179011, rs179009, rs179013, rs179008	SEQUENOM

		rs179012	rs179012	SEQUENOM
		rs179010	rs179010	SEQUENOM
		rs1634319	rs1634319	SEQUENOM
		rs864058	rs850633, rs864058	SEQUENOM
TLR8	X	rs5741883	rs5741883	SEQUENOM
		rs3764880	rs3761624, rs3764879, rs4830805, rs3764880	SEQUENOM
		rs17256081	rs17256081	SEQUENOM
		rs2109134	rs2109134	SEQUENOM
		rs1548731	rs1548731	SEQUENOM
		rs5744067	rs4830808, rs5744067, rs2159377, rs3827469	SEQUENOM
		rs2407992	rs5741886, rs5744069, rs2407992, rs3747414, rs4830807, rs1013151, rs5744080	SEQUENOM
		rs5744088	rs5744055, rs1013150, rs5744088, rs5744068	SEQUENOM

* Genotyping assay failure or not meet the threshold of quality control.

Table S2. Genetic models of inheritance of the risk (A) and the protective (B) haplotypes.

A

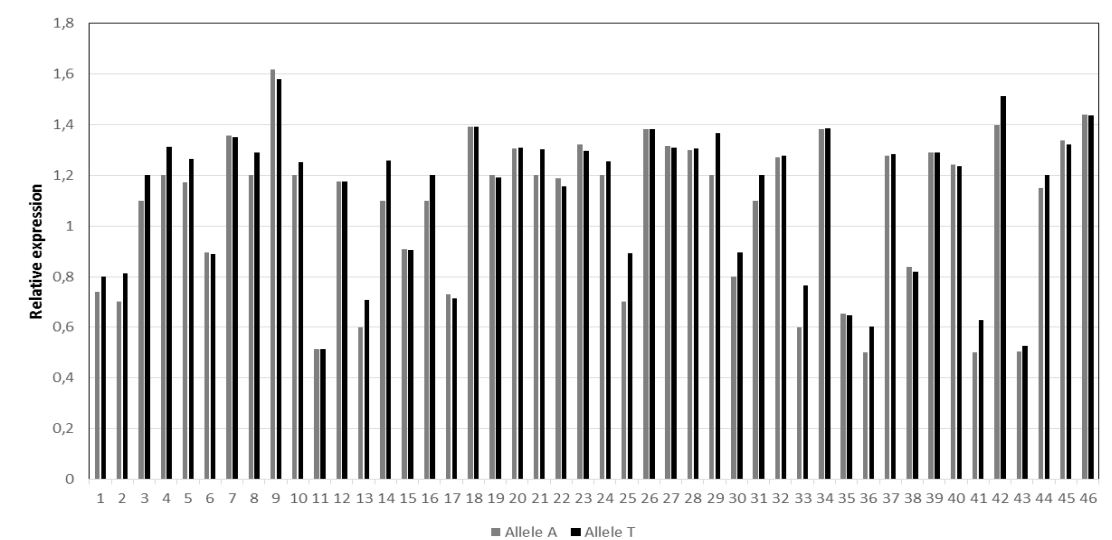
Model	β	Wald Test (p)	AIC	LR (p)
Dominant	0.41577	2.3766 (0.017472)	2349.67	5.5011 (0.019004)
Recessive	0.48185	0.9571 (0.338539)	2354.34	0.8302 (0.362215)
Multiplicative	0.38289	2.4318 (0.015023)	2349.45	5.7241 (0.016734)
General	0.40130	2.2575 (0.023979)	2351.40	5.7731 (0.055768)

B

Model	β	Wald Test (p)	AIC	LR (p)
Dominant	-0.67269	-2.5915 (0.009555)	2346.58	7.5896 (0.005871)
Recessive	-15.3711	-0.0093 (0.992577)	2351.92	3.2560 (0.071163)
Multiplicative	-0.68408	-2.7237 (0.006455)	2346.57	8.6007 (0.003360)
General	-0.62145	-2.4028 (0.016270)	2347.44	9.7301 (0.007712)

β : Estimated haplotype effect. Wald statistic for testing whether $\beta = 0$ and the associated p value. Subjects with 1 copy of the risk haplotype have in the recessive model the same risk as those subjects with no copies; in the dominant model, the same risk as those subjects with 2 copies; in the multiplicative model, an intermediate risk on log scale with respect to those subjects with 0 or 2 copies; and in the general model a general change in risk compared to those with 0 or 2 copies. AIC: Akaike information criteria; LR: likelihood-ratio statistic for testing the global null hypothesis.

Figure S1. Relative expression levels of the alleles A versus T of the single-nucleotide polymorphism nonsynonymous rs6940.



The relative quantification was performed in samples from 46 heterozygous healthy donors. Ratio A/T = 0.97 ± 0.08 ; $p = 0.027$.

Manuscrito en preparación

Large scale genetic analysis in Behçet disease. Identification of residues associated in the HLA Class I region and new susceptibility loci

Lourdes Ortiz-Fernández^{1*}, F. David Carmona^{2*}, Marco-Antonio Montes-Cano¹, José-Raúl García-Lozano¹, Marta Conde-Jaldón¹, Norberto Ortego-Centeno³, María Jesús Castillo⁴, Gerard Espinosa⁵, Genaro Graña-Gil⁶, Juan Sánchez-Bursón⁷, María Rosa Juliá⁸, Roser Solans⁹, Ricardo Blanco¹⁰, Ana-Celia Barnosi-Marín¹¹, R. Gómez de la Torre¹², Patricia Fanlo¹³, Mónica Rodríguez-Carballeira¹⁴, Luis Rodríguez-Rodríguez¹⁵, Teresa Camps¹⁶, Santos Castañeda¹⁷, Juanjo J. Alegre-Sancho¹⁸, Javier Martín², María Francisca González-Escribano^{1*}.

¹Servicio de Inmunología, IBI-S, Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Sevilla; ²Instituto de Parasitología y Biomedicina López Neyra, Consejo Superior de Investigaciones Científicas, Granada. ³Servicio de Medicina Interna, Hospital Clínico San Cecilio, Granada; ⁴Servicio de Medicina Interna, Hospital Universitario Virgen del Rocío, Sevilla; ⁵Servicio de Enfermedades Autoinmunes, Hospital Clinic, Barcelona; ⁶Servicio de Reumatología, Complejo Hospitalario Universitario, A Coruña; ⁷Servicio de Reumatología, Hospital Universitario de Valme, Sevilla; ⁸Servei d'Inmunologia. Hospital Universitari Son Espases. Palma de Mallorca, ⁹Servicio de Medicina Interna, Hospital Vall d'Hebron, Barcelona, ¹⁰Servicio de Reumatología, Hospital Marqués de Valdecilla, Santander; ¹¹Servicio de Medicina Interna, Hospital de Torrecárdenas, Almería, ¹²Servicio de Medicina Interna, Hospital Universitario Central de Asturias, Asturias, ¹³Servicio de Medicina Interna, Hospital Virgen del Camino, Pamplona; ¹⁴Servicio de Medicina Interna, Hospital Universitari Mútua, Terrassa, Terrassa, ¹⁵Servicio de Reumatología, Hospital Clínico San Carlos, Madrid, ¹⁶Servicio de Medicina Interna, Hospital Universitario Carlos Haya, Málaga; ¹⁷Servicio de Reumatología, Hospital de la Princesa, Madrid, ¹⁸Servicio Reumatología, Hospital Universitario Doctor Peset, Valencia.

*These authors contributed equally to this work

Corresponding autor: María Francisca González-Escribano, Servicio de Inmunología. HU Virgen del Rocío. 41013 Sevilla. Spain. mariaf.gonzalez.sspa@juntadeandalucia.es. Tel +34955013228 Fax +34955013221.

Key words: Behçet disease, HLA class I, immunochip, Genetics

INTRODUCTION

Behçet's disease (BD) is an immune-mediated and complex disease characterized by recurrent oral and genital ulceration, ocular involvement, mainly uveitis, and skin lesions, in addition, the disease could affect other parts of body as joints and the central nervous system. The etiology of BD remains poorly understood, although it has been proposed that certain infectious agents and environmental factors are the triggers of disease in genetically predisposed individuals [1]. This disease is most common along the ancient route known as "The Silk Road", stretching from China to the Mediterranean area [2]. This particular geographical distribution in conjunction with the familial aggregation and association with HLA class I molecules (specifically with HLA-B51) are evidences supporting genetic contribution into the pathogenesis of the disease [3,4].

Regarding the genetic component, many studies have described association between this disease and HLA-B51 independently of the ethnicity of the population. Other studies suggest association with other molecules HLA class I (HLA-A, others HLA-B and HLA-C) but these associations are less consistent due primarily to limitations in the size of the cohorts, linkage disequilibrium (LD) in the region and the diversity of typing procedures, which have evolved from serological methods to molecular ones over time. Although, recent studies with high throughput platforms and more large cohorts also support association with others molecules HLA class I, the spectrum of HLA molecules associated with this disease is not completely known [5,6]. It is necessary to clarify this point in order to know what amino acid positions are relevant in the disease predisposition because this information could help to unravel the role of molecules HLA class I in the pathogenesis of the disease.

It has been reported contribution of other genes, non-HLA, to the susceptibility of the disease, but these associations have a less strength and, in general, they are not universally found. There are different genome wide association studies (GWAS) in BD, two in Turkish and three in Asian (Japanese, Chinese and Korean) populations [7-11], although the findings of these studies at GWAS level ($p \leq 10^{-8}$) have been limited because of the number of patients (only one study includes more than 1,000 patients in the discovery cohort), which is necessarily low because BD is a rare disease. Therefore, only by

integrating information from different studies it will be possible to complete the picture of the pathways associated with this disease in different populations. The “Human Immuno DNA Analysis BeadChip Kit” (known as ImmunoChip, Ichip) is a high-density array developed by the ImmunoChip Consortium. This array contains 196,524 variants consisting of single nucleotide polymorphisms (SNPs), rare variants and insertion/deletion (INDEL). These variants are spread across 186 loci which may be involved in susceptibility to autoimmune and inflammatory disorders because of their function [12]. Therefore, this is a high throughput platform with a candidate gene approach, which has been successfully used in many genetic studies in complex diseases such as, rheumatoid arthritis [13], psoriasis [14], systemic sclerosis [15], etc. Additionally, this array is particularly appropriated to extract precise information of the HLA region to use it for imputation because it has the highest coverage of this area, reason it has contributed to the progress in the knowledge of these pathologies both inside as well as outside the HLA region. The aim of this study was to contribute to improve the knowledge of the genetic basis of BD by genotyping a cohort of Spanish patients with this high throughput genotyping platform.

MATERIAL AND METHODS

Study population

This study included a total of 416 BD patient (43.7% males) who fulfilled the 1990 International Study Group classification criteria for Behçet’s disease [16] and 2122 unrelated healthy controls (44.3% males). The discovery cohort consisted of 286 BD patients and 1517 previously genotyped healthy controls and the validation cohort was composed of 130 BD cases and 605 controls. All the subjects were Spanish from European origin and they were recruited from different Spanish hospitals [17]. The study was approved by the local ethics committees of all the participant hospitals and all the individuals included signed a written informed consent.

Genotyping

Peripheral blood or saliva was the starting material obtained from healthy controls and patients. Genomic DNA was extracted using QIAmp DNA Mini Kit (Qiagen,

Barcelona, Spain) according to the manufacturer's recommendations and stored at -20°C. Genotyping in the discovery cohort was performed on the Illumina iScan system using the iChip platform, according to the Illumina protocol

In order to assess the accuracy of the HLA imputation (see below), all the patients from the discovery cohort were also genotyped in the HLA-A, B and C loci using a low resolution (two-digits) PCR-SSOP Luminex method (LABType SSO One Lambda Inc., Canoga Park, CA), according to the manufacturer's instructions. The same HLA typing method was also used to genotype another group of healthy controls (n=377) and to confirm HLA associations in the replication patient cohort (n=130). Replications of the rest of the no-HLA associations detected were performed using TaqMan® SNP Genotyping Assays Assays (Applied Biosystems, Barcelona, Spain) in a 7900HT Fast Real Time PCR System (Applied Biosystems).

Genotyping calling and Quality criteria

Genotype calling was performed after genotyping with the iChip platform according to the manufacturer's instructions with the use of the Illumina iScan System and the Genotyping Module (v.1.8.4) of the GenomeStudio Data Analysis software using the NCBI build 36 (hg18) mapping (Illumina manifest file Immuno_BeadChip_11419691_B.bpm).

Severe Quality Control (QC) parameters were applied in all data set. Those SNPs with a genotyping call rate lower than 90% and those that were not in Hardy-Weinberg equilibrium (HWE; $P < 0.001$) were excluded from analyses. Those samples in which genotyping results were obtained in less than 90% of the SNPs were also excluded. To prevent population stratification, a five first principal-components (PC) analysis was performed by using the "null" SNPs of the iChip. Those samples whose values had standard deviations values greater than 4-fold the centroid of the cluster were considered outlier individuals and were eliminated from analyses. The PC analysis was performed using PLINK [18] and the gcta64 and R-base under GNU Public license v.2. Moreover, when duplicated samples or first-degree relatives were detected, one of the individuals of each duplicated pair or of each pair of relatives were removed using the Genome function with a Pi-HAT threshold 0.5 in PLINK. Of the 286 BD cases of the discovery cohort, 278 passed all these QCs.

Imputation of the HLA region

The SNPs genotyping data of the MHC region at chr6: 20,000,000-40,000,000 bp were used to impute the alleles HLA class I and class II classical (two and four-digit levels) and the polymorphic amino acid positions[19]. The reference panel was comprised of 5,225 individuals of European origin with genotyping data of 8,961 variants (common SNPs and INDEL polymorphisms) across the MHC region together with data of genotyping of the alleles HLA class I and II classical at a level of resolution of four digits [20,21]. The imputation process was performed with the SNP2HLA method using the Beagle software [22]. To impute data, those SNPs included in the iChip platform in the above mentioned region, were filtered according to more stringent QC parameters. Only those variants and those samples with a genotyping success call rate greater than 95% were used for imputation. The final set included a total of 7261 SNPs, 424 alleles HLA classical (two-digits and 4-digits) in the all the HLA class I (HLA-A, HLA-B and HLA-C) and class II (HLA-DRB1, HLA-DQB1, HLA-DQA1, HLA-DPB1 and HLA-DPA1) loci and 1276 of polymorphic amino acid variants (Suppl Table 1). In order to assess the accuracy of the imputation, data of two-digit typing of HLA class-I (HLA-A, HLA-B and HLA-C) obtained by imputation in patients included in the discovery cohort, were compared with those previously obtained in the same individuals using a standard HLA genotyping method (LABType SSO One Lambda Inc., Canoga Park, CA, USA) [22]. Two parameters were used in this comparison: the accuracy, which is a measure of the reliability of the frequencies and a correlation coefficient (r) to establish reproducibility of the typing in each individual.

Statistical analyses

The statistical analyses were performed with PLINK and R. To check association, a logistic regression analysis using the five first PCs as covariates for all the SNPs, all the classical HLA alleles at two-digit and four-digit levels and all the possible combinations of amino acid residues per position was carried out. To identify independent effects, step-wise logistic regression analyses were carried out conditioning by the most associated signals, as described [13]. The omnibus association test was performed to determine the influence of each polymorphic amino acid position in disease susceptibility, as described [22]. A null

generalized linear model, which including the first 5 PCs as variable was built. The alternative model for each position included all the possible amino acids in the analyzed position in addition to the first 5 PCs. Both models, null and alternative, were compared using a likelihood ratio test with $n-1$ degrees of freedom, where n is the number of alternative amino acids at one specific position. Additionally, conditional analyses were conducted by using as conditioning factors the most associated positions. In these conditional analyses, the conditioning factors were included in both models, the null and the alternative. The 3D models of the HLA molecules were performed with the UCSF Chimera software [23].

RESULTS

As expected, the genetic region more strongly associated with the disease was the HLA class I region ($p=2.5 \times 10^{-20}$). Only one additional signal was detected at the GWAS level, rs10889664, corresponding to IL23R ($p=3.8 \times 10^{-12}$), although the p-values for other two SNPs (rs2848479 in the JRKL/CNTN5 region, $p=5.0 \times 10^{-8}$ and rs1874886 in IL12A, $p=6.7 \times 10^{-8}$) were very close to this threshold. In addition, another 66 signals were detected in the grey area (p-values from 10^{-5} to 10^{-8}) (Figure 1).

Regarding the HLA region, a high overall accuracy of the imputation was observed for the three loci HLA-class I included: HLA-A, HLA-B and HLA-C when comparing imputation results with those obtained for same individuals using a standard HLA genotyping method (Suppl Table 2 and Figure S1). Additionally, the correlation coefficient (r) was also adequate for all the allele-groups except for HLA-A*25 and HLA-A*26 (Suppl Table 2). As a result of the imputation of the HLA region, a total of 7,227 SNPs, 1,273 polymorphic amino acid variants and 419 classical allele-groups HLA class I and II (125 at two-digit and 294 at four-digit resolution levels) which had passed all QCs were included in the analysis. Frequencies obtained for all the classic allele-groups HLA class I and II at two and four digits resolution after imputation are displayed in Suppl Tables 3 and 4. After performing a logistic regression analysis including the 5PCs as covariates in order to control for population stratification, multiple signals of association at GWAS level were detected. As expected, the highest association peaks were observed in the HLA-B region,

although several signals throughout all the class I subregions were detected, however, no signal was detected in any of the different HLA class II subregions (Suppl Tables 5). To avoid the strong LD in the HLA region, a stepwise conditional analysis was performed in order to identify those HLA class I subregions that independently influence the susceptibility to BD. Our results confirm an independent association of variants of both HLA-A (A*02) and HLA-B (B*51, B*57 and B*15), nevertheless, association of HLA-C was completely abolished by introduction of the major associated variants as covariates (Table 1). Next, to replicate association on the HLA-A and B subregions, another set of patients were genotyped. Several HLA-A and B allele-groups were found associated in the overall cohort (Suppl Table 6). The multivariate analysis revealed independent association of A*03, B*15, B*51 and B*57 (Suppl Table 7). When individuals were stratified according to their B*51 or B*15+51+57 status, frequencies of A*02 and A*03 had the same trend in all the groups (Suppl Table 8). Next, association of the polymorphic amino acid residues in all the HLA molecules was investigated by using the omnibus test (Suppl Table 8). According to this analysis, the amino acid position strongest associated with susceptibility to BD is the position 97 in HLA-B. Thr and Val confer risk at this position (OR=2.74 and 2.71, respectively) whereas Arg and Ser confer protection (OR= 0.57 and 0.72, respectively). Regarding HLA-A, the most associated position is the 66 being Lys the amino acid conferring risk (OR=1.64), this association remains significant after conditioning by position 97 in HLA-B (Figures 2 and 3). Practically identical results were obtained by performing a stepwise conditional analysis including only the amino acid residues of HLA-A and HLA-B (Suppl Table 10). Table 2 provides a summary of the classical HLA alleles encoding amino acids of interest in relevant positions and Figure 4 displays the location of these positions in the 3D structure of the HLA class I molecule.

For the rest of the loci, only 3 SNPs were found associated at GWAS level or very close to it, although other 66 were found in the grey area (Table Suppl 11). With regards to the GWAS level, two SNPs were selected for replication: rs2848479 (between *JFK* and *CNTN5*) and rs1874886 (*IL12A*). In addition, another SNP in *IL12A*, rs17810546, was genotyped in the entire cohort because it had been found associated in a previous study in another population. *IL23R* was not included because it has been described as associated with susceptibility to BD in different populations including ours. From all the genes

included in the grey area, *STAT4*, *CCR1* and *UBASH3B* were selected for replication since they have been associated with this disease in previous studies in other populations. Furthermore, *KIR2DL4* was included because it is a very particular KIR molecule both structurally and functionally and it has been related with severe uveitis in BD. In the cases of rs2848479 (*JRKL/CNTN5*), rs1874886 and rs17810546 (*IL12A*), rs11694530 (*STAT4*) and rs7616215 (*CCR1*), significance was also reached in the replication cohort. For rs604999 (*KIR2DL4*) and rs7944144 (*UBASH3B*) the differences were not significant in the replication cohort but the trend was the same as in the discovery cohort. In all the cases, the differences were significant in the overall cohort (Table 3).

DISCUSSION

The first conclusion of this study which was performed with a high-density array, iChip, is that two of the subregions of HLA class I, HLA-A and HLA-B, are independently associated to BD. Also, our analysis discards an association with the susceptibility to this disease of the other region of class I, HLA-C and of the entire class II region. In this sense, our results are in agreement with those recently obtained by Ombrello et al in the Turkish [5] population using a different array and other population. As in the study by Ombrello et al, our results demonstrated, the importance of the locus HLA-B itself in the susceptibility to this pathology, because, no evidences of independent association with rs116799036 or with other SNP in the B-MICA region were found. Our data, support that the association between SNPs in the HLA-B-MICA region and the disease is a consequence of LD of markers located in this region with B*51. Therefore, in this respect, our results are contrary to those obtained by Hughes et al [6], suggesting that this noncoding SNP located between HLA-B and MICA was the true causal variant of the association of HLA-B*51 BD. Moreover our results also demonstrated association of several HLA-A and HLA-B allele-groups with this disease.

In the present study, a severe quality control of the imputation of the three loci of HLA class I was performed in order to validate the method and the reference panel used in imputation. A relatively large subset of individuals was genotyped by using both methods, traditional HLA typing as well as imputation, being values of accuracy appropriate for all

the allele-groups. (>95%). Nevertheless, inappropriate r-values were detected in the cases of A*25 and A*26 because imputation failed in some of these individuals by assigning as A*25, individuals genotyping as A*26 by traditional methods and vice versa. These wrong assignments could be caused by the set of SNP that was used to impute these allele-groups, which could be not completely appropriate or because the data base of individuals used as reference for imputation would contain some errors. In any case, results of this parameter, which is not included in most of the studies, invalidate, in the cases of A*25 and A*26, the comparison between groups in which HLA typing has been obtained using traditional methods with those in which this typing has been imputed. In addition, the low r-values do not permit the inclusion of the data of these two allele-groups, obtained by imputation, in subsequent analyses involving stratification. This limitation could be related with discrepancies observed between different cohort from the same population.

Several HLA class I allele-groups are associated with the disease in the discovery cohort, nevertheless, the high LD degree in a region that includes many important genes related with the immune response complicates the interpretation of the results. As a result of the multivariate analysis, association with the HLA-C region was completely abolished. In this respect, our results are also similar to Ombrello et al [5] and differ from Hughes et al [6]. According to these analyses, B*51 was the strongest associated factor but other allele-groups HLA-B are independently associated with the disease, specifically, B*15, B*51 and B*57. Therefore, the stepwise in the discovery cohort and the multivariate analysis in the overall reveal the same results supporting at least 3 risk factors in the locus B. No significant differences were found in the case of other allele groups such as HLA-B*27, B*49, and others, in which trend towards association has been described in some populations. Although the association with disease can not be completely discarded in these cases, the force of the association would be slight because frequency of HLA-B*27 and B*49 in our population is higher than those of other allelic groups, such as B*57, that have been detected associated. Several allele-groups HLA-B, specifically, B*18, B*35, B*38 and B*58 were found associated as protective factors in the overall cohort, nevertheless, none remains significant in multivariate analyses. Regarding HLA-A, our results support independent association of this locus, but results in allele-associated groups are less consistent because according to the stepwise analysis, A*02 is the risk factor that remain

significant after controlling by B*51+57+15 whereas, A*03 is the HLA factor (protective) that remain significant, after multivariable analysis in the overall cohort. A*02, is in a strong LD with B*51, but frequency of HLA-A*02 was higher all the groups of patients (B*51 positive and negative and B15-51+57 positive and negative) compared with counterpart controls, although significance was reached only in those groups without HLA-B risk factors (B*51 negative and B15+51+57 negative). These results suggest independent association of A*02, although the multivariate analysis does not reach significance. Regarding A*03, its frequency was lower in all the patient groups and the association was detected also in the multivariate but not in the stepwise analysis. Regarding other HLA allele-groups, some studies have suggested that A*26 is a risk factor for the disease. HLA-A*26 was eliminated of the analysis performed with the discovery cohort because of the low r-value obtained by comparing the results of the imputation with traditional typing for this allele-group. However, the distribution of A*26 was not significantly different when comparing patients and controls genotyped using a standard method for HLA, moreover, the trend was toward a lower frequency in patients than in controls (0.03 vs. 0.06). Lastly, in a previous study, our group described independent association of A*24 with BD, but this association was not confirmed in the present study.

Results in the HLA class I region illustrate the difficulty to evaluate association of the different subregions. This is particularly important when, as is the case of the study of the classical HLA alleles, the statistical power for each allele is different depending on its frequency in the population under study. Not always the most significant factors in univariate analyses are independently associated (as is the case of HLA-C) whereas other factors with P-values with lesser statistical significance are independently associated (for example, A*03 and B*15). Finally, it begins to be clear that there are several allelic groups class I are associated with BD and it is necessary to establish precisely what these factors are in order to generate hypotheses about the etiopathogenic role of these molecules in this disease. The study of amino acids confirms the relevance of a set of positions of the HLA-B molecules in the disease. Among these positions, 97B is the most strongly associated with the disease according to the Omnibus test. The p-values for the top HLA-C positions were dramatically increased when the test was conditioned to 97B, indicating dependency, ie LD, nevertheless, p-values for the top the set of HLA-A aminoacid remain approximately

equal indicating independency. On the other hand, other HLA-B top amino acid positions showed dependency of the 97B, nevertheless, from our point of view, the biological interpretation is different in this case. It is evident that each specific HLA-B molecule has a specific set of amino acid in polymorphic positions which, as a consequence, they are in LD and they are dependent but dependency in this case does not excluded influence of this position. In this sense, these kind of studies discarded the irrelevant positions and give an idea of which may be the most relevant but only by complementing this knowledge with those provided by functional studies would permit clarify this question. It has been demonstrated interaction of a inhibitory receptor of the NK cells (KIR3DL1) and molecules HLA with a serological-defined epitope named Bw4 (determined by amino acid positions 79–83 of the molecule) having Ile80 (Bw4-I80). This interaction has an inhibitory effect on the cytotoxic capacity of NK cells. Also, it has been described that changes at specific positions of the HLA-B outside the Bw4 epitope, affect the interaction of Bw4 with KIR3DL1. Therefore results of this test additionally support independent association of the loci A and B and lack of association of the HLA and the class II molecules.

Regarding the rest of the genes, none previous studies using Immunochhip as platform has been reported data. Association between IL-23R and this disease has been reported in different populations, included Spanish [8-9,24]. The specific position associated with the disease remains unknown and the SNP described are just marked but the results highlighted the importance of IL23R in the susceptibility to BD. The rs2848479 is located in an intergenic region between the human homolog of mouse jerky gene (HHMJG, JRKL) and the contactin 5 (CTCN5) genes. JRKL encodes a protein with an unknown function and CTCN5 encoding a member of the contactin family (member belongings to the immunoglobulin superfamily) which mediate cell surface interactions during nervous system development. IL12A, is a subunit of the heterodimeric cytokines: IL12 and IL35 [25]. It has been proposed as suggestive of association with BD in a GWAS in the Turkish population an in a meta-analysis including this data together those obtained of a mixed population [26]. The SNP associated in the GWAS was rs17810546 which is not in LD with rs1874886 detected in discovery cohort of our study. In the replication both SNPs were included and association was confirmed for these 2 SNPs. Therefore, results support association of this gene in BD disease in other population.

In conclusion, our study, performed with HLA classical typing, supports that the association of HLA-B*51 with Behçet disease is explained by the HLA-B itself better than. by a variant located between HLA-B and MICA This study supports two different loci of susceptibility independently associated with the disease in the HLA region, specifically, HLA-A and HLA-B. Nevertheless, our results do not support independent association of the HLA-C region.

REFERENCES

1. Mendes D, Correia M, Barbedo M, Vaio T, Mota M, Gonçalves O, Valente J. Behçet's disease: a contemporary review. *J Autoimmun* 2009; 32:178-88.
2. Sakane T, Takeno M, Suzuki N, Inaba G. Behçet's disease. *N Engl J Med* 1999, 341:1284-1291.
3. Koné-Paut I, Geisler I, Wechsler B, Ozen S, Ozdogan H, Rozenbaum M, Touitou I. Familial aggregation in Behçet's disease: high frequency in siblings and parents of pediatric probands. *J Pediatr* 1999; 135:89-93.
4. Menthon de M, Lavalley MP, Maldini C, Guillevin L, Mahr A. HLA-B51/B5 and the risk of Behçet's disease: a systematic review and meta-analysis of case-control genetic association studies. *Arthritis Rheum* 2009; 61:1287-96.
5. Ombrello MJ, Kirino Y, de Bakker PI, Gül A, Kastner DL, Remmers EF. Behçet disease-associated MHC class I residues implicate antigen binding and regulation of cell-mediated cytotoxicity. *Proc Natl Acad Sci USA*. 2014; 111 (24):8867-72.
6. Hughes T, Coir P, Alder A, Yilmaz V, Aksu K, Düzgün N, et al. Identification of multiple independent susceptibility loci in the HLA region in Behçet's disease. *Nat Genet*. 2013; 45 (3):319-24.
7. Fei Y, Webb R, Cobb BI, Direskeneli H, Saruhan-Direskeneli G, Sawalha AH. Identification of novel genetic susceptibility loci for Behçet's disease using a genome-wide association study. *Arthritis Res Ther*. 2009; 11 (3):R66.
8. Remmers EF, Cosan F, Kirino Y, Ombrello MJ, Abaci N, Satorius C, et al. Genome-wide association study identifies variants in the MHC class I, IL10, and IL23R-IL12RB2 regions associated with Behçet's disease. *Nat Genet*. 2010; 42 (8):698-702.

9. Mizuki N, Meguro A, Ota M, Ohno S, Shiota T, Kawagoe T, et al. Genome-wide association study identify IL23R-IL12RB2 and IL10 as Behçet's disease susceptibility loci. *Nat Genet.* 2010; 42 (8):703-6.
10. Hou S, Yang Z, Du L, Jiang Z, Shu Q, Chen Y, et al. Identification of a susceptibility locus in STAT4 for Behçet's disease in Han Chinese in a Genome-wide association study. *Arthritis Rheum.* 2012; 64 (12):4104-13.
11. Lee YJ, Horie Y, Wallace GR, Choi YS, Park JA, Choi JY, et al. Genome-wide association study identifies GIMAP as a novel susceptibility locus for Behçet's disease. *Ann Rheum Dis.* 2013; 72 (9): 1510-1516.
12. Cortes A, Brown MA. Promise and pitfalls of the Immunochip. *Arthritis Res Ther.* 2011; 13 (1):101.
13. Eyre S, Bowes J, Diogo D, Lee A, Barton A, Martin P, *et al.* High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. *Nat Genet.* 44, 1336-1340.
14. Tsoi LC, Spain SL, Knight J, Ellinghaus E, Stuart PE, Capon F, *et al.* Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity. *Nat Genet.* 44, 1341-1348.
15. Mayes MD, Bossini-Castillo L, Gorlova O, Martin JE, Zhou X, Chen WV, *et al.* Immunochip analysis identifies multiple susceptibility loci for systemic sclerosis. *Am J Hum Genet* 94, 47-61.
16. International Study group for Behcet's disease. Criteria for diagnosis of Behcet's disease. *Lancet.* 1990; 335:1078-1080.
17. Carmona FD, Mackie SL, Martín JE, Taylor JC, Vaglio A, Eyre S, et al. A large-scale genetic analysis reveals a strong contribution of the HLA class II region to giant cell arteritis susceptibility. *Am J Hum Genet.* 2015; 96 (4): 565-80.
18. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferrerira MA, Bender D, *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007; 81: 559-575.
19. Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodiyar VK, *et al.* Gene map of the extended human MHC. *Nat Rev Genet.* 2004; 5, 889-899.

20. Brown WM, Pierce J, Hilner JE, Perdue LH, Lohman K, Li L, *et al.* Overview of the MHC fine mapping data. *Diabetes Obes Metab* 2009; 11: 2,7.
21. de Bakker PI, McVean G, Sabeti PC, Miretti MM, Green T, Marchini J, *et al.* A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat Genet.* 2006; 38:1166-1172.
22. Raychauduri S, Sandor C, Stahl EA, Freudenberg J, Lee HS, JiaX, *et al.* Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. *Nat Genet.* 2012; 44, 291-296.
23. Pettersen EF, Goddard TD, Haung CC, Couch GS, Greenblatt DM, Meng EC, *et al.* UCSF Chimera- a visualization system for exploratory research and analysis. *J Comput Chem.* 2004, 25, 1605,1612.
24. Montes-Cano MA, Conde-Jaldón M, García-Lozano JR, Ortiz-Fernández L, Ortego-Centeno N, Castillo-Palma MJ, *et al.* HLA and non-HLA genes in Behçet's disease: a multicentric study in the Spanish population. *Arthritis Res Ther.* 2013, 15 (5): R145.
25. Chang JT, Shevach EM, Segal BM. Regulation of interleukin IL12 receptor beta2 subunit expression by endogenous IL12: a critical step in the differentiation of pathogenic autoreactive T cells. *J Exp Med.* 1999; 189 (6):969e978.
26. Kirino Y, Bertsias G, Ishigatsubo Y, Mizuki N, Tugal-Tutkun I, Seyahi E, *et al.* Genome-wide association analysis identifies new susceptibility loci for Behçet's disease and epistasis between HLA-B*51 and ERAP1. *Nat Genet.* 2013; 45 (2):202-207.

Table 1. Stepwise analysis to detect independent associations in the class I region

Covariates	Risk Allele	Pvalue	OR (95%CI)
None	HLA-B*51	1.182e-26	3.63 (2.87-4.60)
HLA-B*51	HLA-B*57	1.058e-6	2.75 (1.83-4.12)
HLA-B*51+B*57	HLA-B*15	4.68e-05	2.10 (1.47-3.01)
HLA-B*51+B*57+B*15	HLA-A*02	0.002	1.36 (1.12-1.64)

+ = HLA-A*25 and HLA-A*26 are not included in this analysis because of the correlation coefficient (r) was low.

Table 2. Classical HLA allele-groups having amino acid residues influencing the risk to BD in the most relevant positions.

Molecule	Amino acid position	Residue	Residue frequency		P-value	OR [CI 95%]	Classical HLA alleles
			BD	CTRL			
HLA-B	97	Thr+Val	0.344	0.165	2.32E-21	2.86(2.31-3.56)	HLA-B*13:02, HLA-B*51, HLA-B*52, HLA-B*54, HLA-B*55, HLA-B*56, HLA-B*57, HLA-B*59, HLA-B*78
		Arg+Ser	0.565	0.744	1.97E-16	0.45 (0.37-0.54)	HLA-B*07, HLA-B*08, HLA-B*15, HLA-B*18, HLA-B*35, HLA-B*37, HLA-B*38, HLA-B*39, HLA-B*41, HLA-B*42, HLA-B*44, HLA-B*46, HLA-B*48, HLA-B*49, HLA-B*53, HLA-B*58, HLA-B*81, HLA-B*82
HLA-A	66	Lys	0.493	0.368	1.29E-06	1.58(1.32-1.91)	HLA-A*02, HLA-A*23, HLA-A*24, HLA-A*34

The table shows those groups of HLA alleles in which the most common alleles have the amino acid of interest in these positions

Table 3. Replication study of non-HLA genes

Chr	Gene	Variant	Allele	Ichip				Taqman				Combined			
				BD (n=278)	Controls (n=1517)	P-v	OR (CI 95%)	BD (n=130)	Controls (n=605)	P-v	OR (CI 95%)	BD (n=408)	Controls (n=2122)	P-v	OR (CI 95%)
11	JRKL/ CNTN5	rs2848479	A	0.529	0.404	5.00E-08	1.65 (1.38-1.98)	0.505	0.391	0.002	1.59 (0.17-2.15)	0.523	0.401	7.28E-10	1.64 (1.40-1.92)
3	IL12A	rs1874886	A	0.492	0.367	6.67E-08	1.67 (1.39-2.02)	0.447	0.378	0.039	1.40 (1.02-1.94)	0.478	0.370	1.28E-08	1.60 (1.36-1.88)
3	IL12A	rs17810546	G	0.146	0.098	6.67E-04	1.58 (1.21-2.06)	0.159	0.084	0.002	2.07 (1.29-3.33)	0.149	0.094	9.23E-06	1.68 (1.33-2.11)
2	STAT4	rs11694530	A	0.520	0.399	7.44E-07	1.62 (1.35-1.96)	0.482	0.378	0.013	1.53 (1.09-2.14)	0.499	0.394	1.86E-07	1.53 (1.30-1.79)
19	KIR2DL4	rs604999	A	0.215	0.144	7.75E-05	1.63 (1.30-2.04)	0.161	0.141	ns		0.200	0.143	6.03E-05	1.49 (1.23-1.82)
11	UBASH3B	rs7944144	G	0.360	0.455	5.76E-05	0.67 (0.56-0.82)	0.401	0.442	ns		0.378	0.450	3.00E-04	0.74 (0.63-0.88)
3	CCR1	rs7616215	C	0.306	0.347	0.057	0.83 (0.68-1.01)	0.259	0.353	0.016	0.64 (0.44-0.92)	0.294	0.349	0.004	0.78 (0.66-0.93)

Figure Legends

Figure1. Manhattan plot representation of the Immunochip results. The $-\log_{10}$ of the combined logistic regression test P-values are plotted against its physical chromosomal position. The red line represents the genome-wide level of significance ($P < 5 \times 10^{-8}$). The blue line represents the statistical significance for suggestive signal ($P < 1 \times 10^{-4}$). The most relevant associations are highlighted.

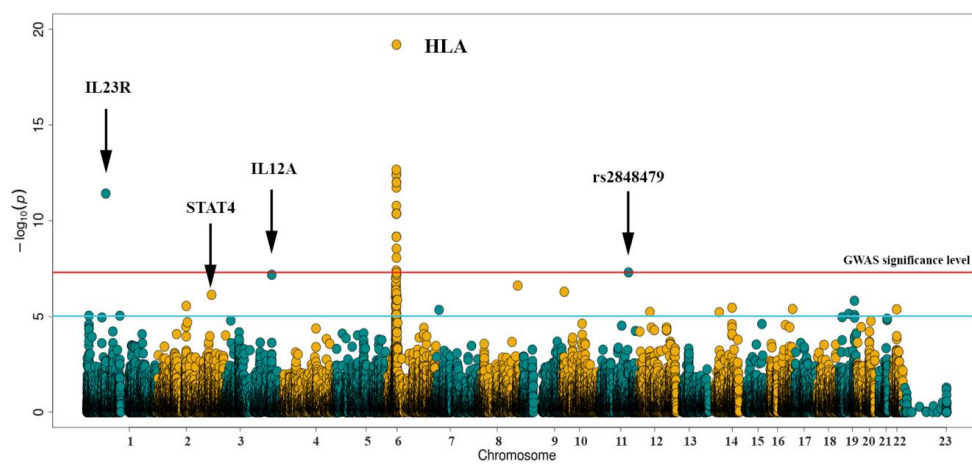
Figure2. Manhattan plot representation of the Omnibus test of the amino acid positions. A) Unconditioned test. B) Results of Omnibus test after conditioning to HLA-B position 97. C) Results of Omnibus test after conditioning to HLA-B position 97 and HLA-A position 66. The $-\log_{10}$ of the combined logistic regression test P-values are plotted against its physical chromosomal position. The red line represents the genome-wide level of significance ($P < 5 \times 10^{-8}$).

Figure3. Independent analysis of the different amino acids of HLA-B position 97 and HLA-A position 66. Each row refers to a single amino acid. Allele frequencies of cases and controls, odd ratios (OR) and $-\log_{10}$ of the logistic regression test P-values are plotted.

Figure4. Ribbon representation of the HLA molecules HLA-B and HLA-A. The associated amino of the HLA model associated with BD are highlighted in red.

Supplementary Figure1. Imputation accuracy in this study of classical HLA-A, HLA-B and HLA-C alleles.

Figure 1.



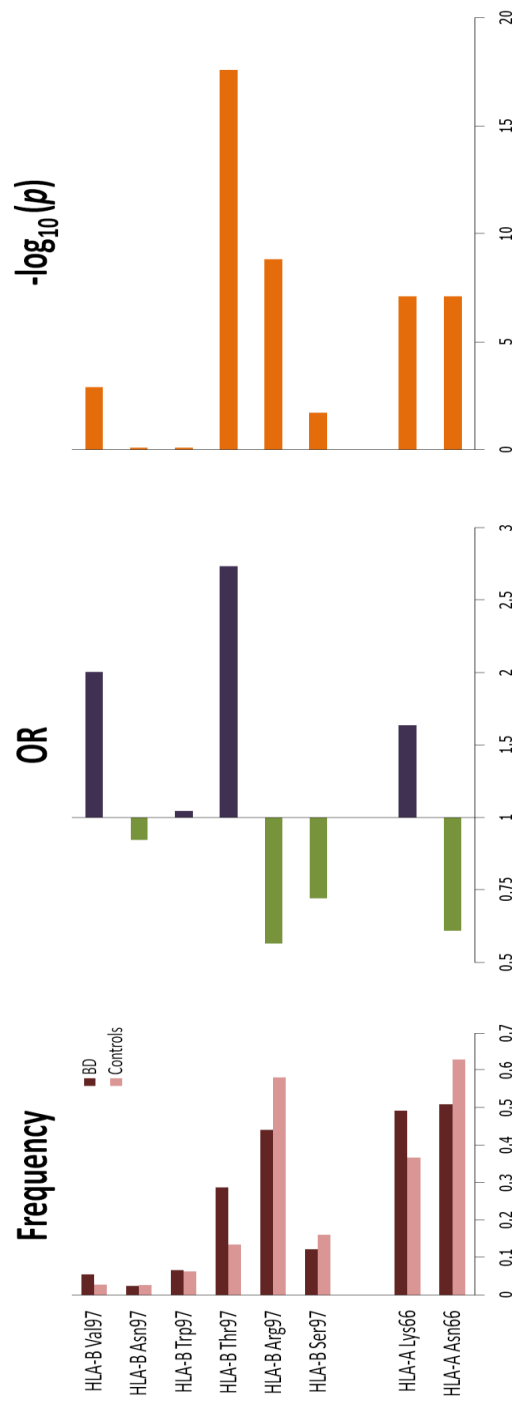


Figure 3.

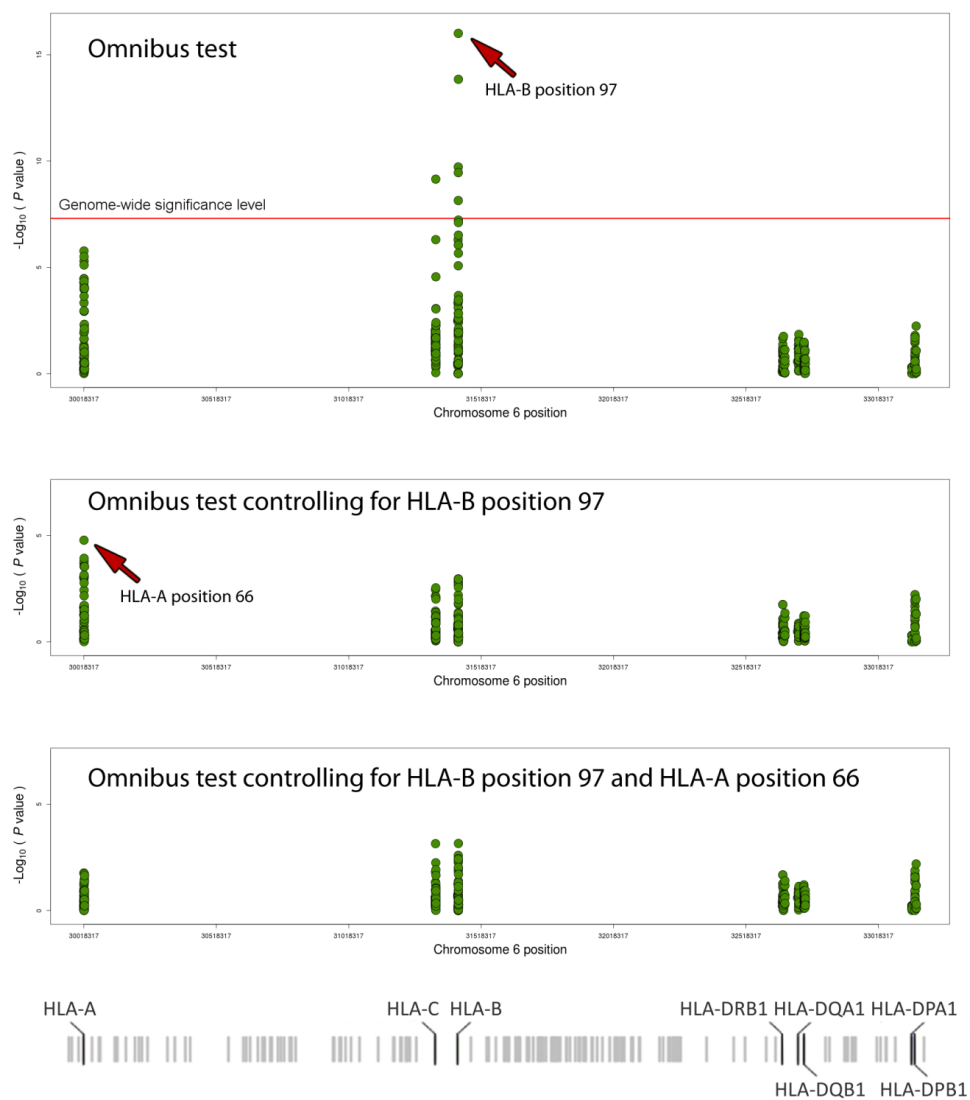


Figure 4.

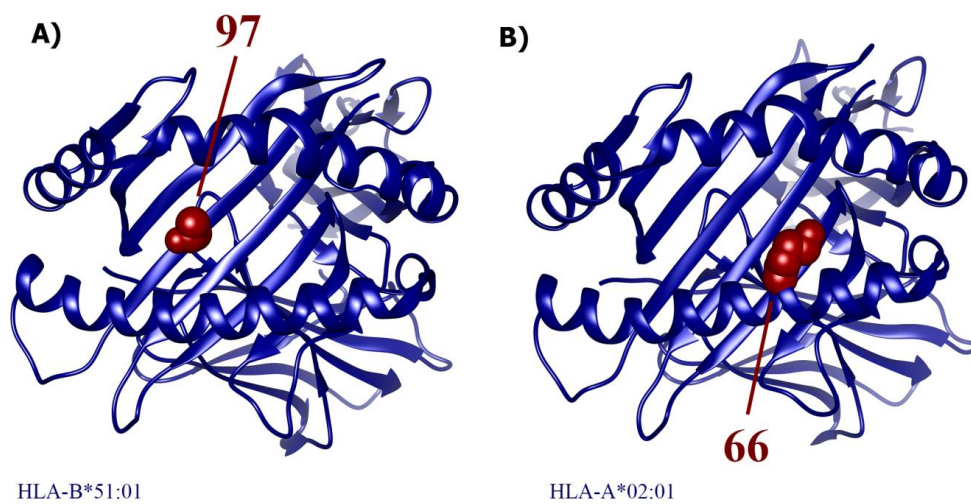
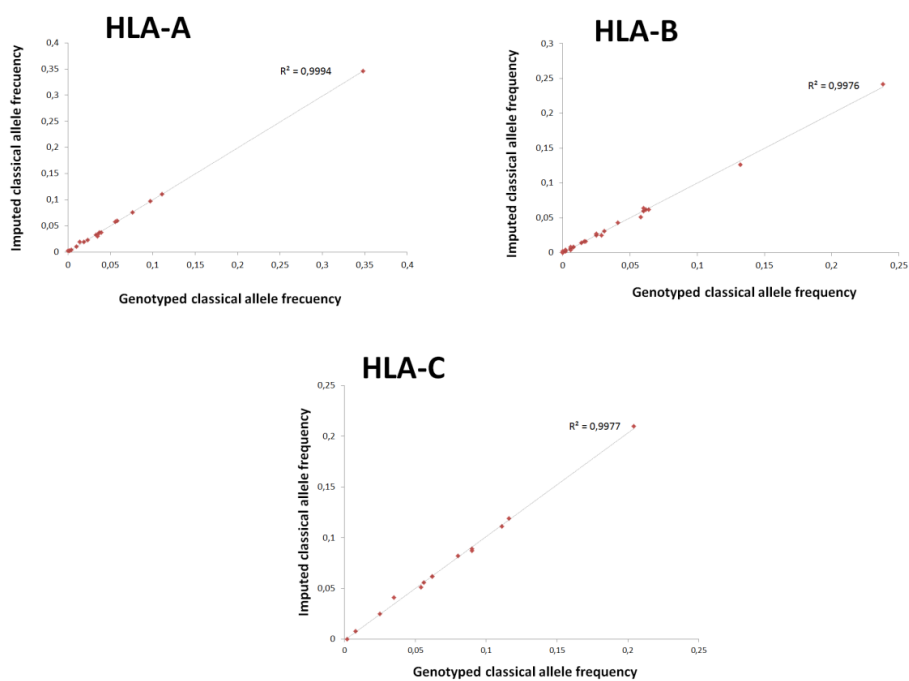


Figure S1



Suppl. Table 3. Univariate analysis of the classical allele groups in the HLA region at two-digit resolution level

Gene	Allele	Allele Frequency		P-value	OR (95%CI)
		BD (n=278)	Controls (n=1517)		
HLA-A	HLA-A*01	0.103	0.123	ns	
	HLA-A*02	0.344	0.247	3.09E-06	1.60 (1.32-1.94)
	HLA-A*03	0.063	0.114	0.0004	0.52 (0.36-0.75)
	HLA-A*11	0.059	0.064	ns	
	HLA-A*23	0.038	0.028	ns	
	HLA-A*24	0.11	0.093	ns	
	HLA-A*29	0.072	0.08	ns	
	HLA-A*30	0.034	0.056	0.037	0.60 (0.37-0.97)
	HLA-A*31	0.032	0.022	ns	
	HLA-A*32	0.034	0.039	ns	
	HLA-A*33	0.018	0.026	ns	
	HLA-A*34	0.004	0.002	ns	
	HLA-A*36	0.002	0	ns	
	HLA-A*66	0.009	0.006	ns	
	HLA-A*68	0.002	0.032	ns	
	HLA-A*69	0.002	0.006	ns	
	HLA-A*74	0.002	0.001	ns	
	HLA-A*80	0.002	0.002	ns	
HLA-B	HLA-B*07	0.065	0.082	ns	
	HLA-B*08	0.045	0.063	ns	
	HLA-B*13	0.009	0.016	ns	
	HLA-B*14	0.067	0.064	ns	
	HLA-B*15	0.061	0.047	ns	
	HLA-B*18	0.058	0.082	0.046	0.68 (0.47-0.99)
	HLA-B*27	0.025	0.027	ns	
	HLA-B*35	0.058	0.099	0.002	0.55 (0.38-0.81)
	HLA-B*37	0.004	0.007	ns	
	HLA-B*38	0.023	0.036	ns	
	HLA-B*39	0.013	0.016	ns	
	HLA-B*40	0.023	0.033	ns	
	HLA-B*41	0.005	0.017	ns	
	HLA-B*42	0.002	0.001	ns	
	HLA-B*44	0.124	0.16	0.031	0.74 (0.57-0.98)
	HLA-B*45	0.016	0.015	ns	
	HLA-B*46	0.002	0.001	ns	
	HLA-B*47	0.002	0.001	ns	
	HLA-B*48	0	0.001	-	
	HLA-B*49	0.036	0.04	ns	
	HLA-B*50	0.025	0.025	ns	
	HLA-B*51	0.245	0.08	7.9E-27	3.72 (2.94-4.70)
	HLA-B*52	0.007	0.015	ns	

	HLA-B*53	0.014	0.011	ns	
	HLA-B*54	0	0	-	
	HLA-B*55	0.007	0.014	ns	
	HLA-B*56	0.004	0.004	ns	
	HLA-B*57	0.058	0.028	0.001	2.01 (1.31-3.05)
	HLA-B*58	0.005	0.014	ns	
	HLA-B*73	0	0.001	-	
	HLA-B*78	0	0	-	
	HLA-B*81	0	0	-	
HLA-C	HLA-C*01	0.025	0.034	ns	
	HLA-C*02	0.063	0.04	0.019	1.60 (1.09-2.36)
	HLA-C*03	0.05	0.067	ns	
	HLA-C*04	0.094	0.128	0.022	0.70 (0.52-0.95)
	HLA-C*05	0.081	0.107	ns	
	HLA-C*06	0.083	0.065	ns	
	HLA-C*07	0.21	0.254	0.03	0.78 (0.63-0.97)
	HLA-C*08	0.068	0.061	ns	
	HLA-C*12	0.047	0.083	0.004	0.54 (0.36-0.82)
	HLA-C*14	0.049	0.018	2.34E-05	2.82 (1.76-4.51)
	HLA-C*15	0.103	0.041	3.45E-09	2.70 (1.95-3.75)
	HLA-C*16	0.119	0.085	0.01	1.44 (1.08-1.92)
	HLA-C*17	0.007	0.016	ns	
	HLA-C*18	0.002	0.001	ns	
HLA-DQ	HLA-DQA1*01	0.340	0.377	0.098	
	HLA-DQA1*02	0.187	0.166	ns	
	HLA-DQA1*03	0.166	0.138	0.094	
	HLA-DQA1*04	0.045	0.025	0.011	1.80 (1.14-2.83)
	HLA-DQA1*05	0.259	0.293	0.1033	
	HLA-DQA1*06	0.005	0.003	ns	
	HLA-DQB1*02	0.268	0.278	ns	
	HLA-DQB1*03	0.345	0.317	ns	
	HLA-DQB1*04	0.049	0.030	0.026	1.64 (1.06-2.52)
	HLA-DQB1*05	0.166	0.183	ns	
	HLA-DQB1*06	0.173	0.194	ns	
	HLA-DPA1*01	0.228	0.238	ns	
	HLA-DPA1*02	0.227	0.236	ns	
	HLA-DPA1*03	0.002	0.003	ns	
	HLA-DPA1*04	0	0	ns	
	HLA-DPB1*01	0.043	0.047	ns	
	HLA-DPB1*02	0.196	0.168	ns	
	HLA-DPB1*03	0.081	0.085	ns	
	HLA-DPB1*04	0.478	0.483	ns	
	HLA-DPB1*05	0.022	0.017	ns	
	HLA-DPB1*06	0.009	0.015	ns	
	HLA-DPB1*09	0	0.005	ns	
	HLA-DPB1*10	0.020	0.024	ns	
	HLA-DPB1*11	0.045	0.043	ns	

HLA-DP	HLA-DPB1*13	0.043	0.025	0.018	1.78(1.11-2.88)
	HLA-DPB1*14	0.007	0.031	0.003	0.22(0.08-0.60)
	HLA-DPB1*15	0.007	0.009	ns	
	HLA-DPB1*16	0.005	0.004	ns	
	HLA-DPB1*17	0.036	0.035	ns	
	HLA-DPB1*18	0	0	ns	
	HLA-DPB1*19	0.004	0.004	ns	
	HLA-DPB1*20	0	0	ns	
	HLA-DPB1*21	0	0	ns	
	HLA-DPB1*22	0	0	ns	
	HLA-DPB1*23	0.002	0.002	ns	
	HLA-DPB1*26	0.002	0.001	ns	
	HLA-DPB1*28	0	0	ns	
	HLA-DPB1*30	0	0	ns	
	HLA-DPB1*31	0	0	ns	
	HLA-DPB1*33	0	0	ns	
	HLA-DPB1*34	0	0	ns	
	HLA-DPB1*35	0	0	ns	
	HLA-DPB1*40	0	0.001	ns	
	HLA-DPB1*45	0	0	ns	
	HLA-DPB1*48	0	0	ns	
	HLA-DPB1*81	0	0	ns	
	HLA-DPB1*85	0	0	ns	
HLA-DR	HLA-DRB1*01	0.097	0.115	ns	
	HLA-DRB1*03	0.103	0.132	0.06	
	HLA-DRB1*04	0.155	0.128	0.09	
	HLA-DRB1*07	0.187	0.165	ns	
	HLA-DRB1*08	0.045	0.028	0.036	1.63(1.03-2.57)
	HLA-DRB1*09	0.009	0.010	ns	
	HLA-DRB1*10	0.022	0.016	ns	
	HLA-DRB1*11	0.153	0.139	ns	
	HLA-DRB1*12	0.007	0.011	ns	
	HLA-DRB1*13	0.106	0.124	ns	
	HLA-DRB1*14	0.025	0.028	ns	
	HLA-DRB1*15	0.077	0.091	ns	
	HLA-DRB1*16	0.014	0.018	ns	

Suppl. Table 4. Univariate analysis of the classical allele groups in the HLA region at four-digit resolution level

Allele	Allele Frequency		P-value	OR (95%CI)
	BD (n=278)	Controls (n=1517)		
HLA-A*0101	0.103	0.119	ns	1.65 (1.35-2.01)
HLA-A*0102	0	0.004	ns	
HLA-A*0103	0	0.001	ns	
HLA-A*0201	0.323	0.223	7.71E-07	
HLA-A*0202	0.002	0.006	ns	
HLA-A*0203	0	0.001	ns	
HLA-A*0205	0.018	0.017	ns	
HLA-A*0206	0.002	0.001	ns	
HLA-A*0207	0	0	-	
HLA-A*0211	0	0	-	
HLA-A*0216	0	0	-	0.53 (0.37-0.76)
HLA-A*0217	0	0	-	
HLA-A*0301	0.06115	0.1111	0.0006	
HLA-A*0302	0.002	0.003	ns	
HLA-A*1101	0.059	0.064	ns	
HLA-A*1102	0	0	-	
HLA-A*1103	0	0	-	
HLA-A*1105	0	0	-	
HLA-A*2301	0.038	0.028	ns	
HLA-A*2401	0	0	-	3.46 (1.12-10.66)
HLA-A*2402	0.110	0.093	ns	
HLA-A*2403	0	0	-	
HLA-A*2407	0	0	-	
HLA-A*2410	0	0	-	
HLA-A*2901	0.009	0.003	0.03	
HLA-A*2902	0.063	0.078	ns	
HLA-A*3001	0.005	0.011	ns	0.54 (0.31-0.97)
HLA-A*3002	0.023	0.042	0.035	
HLA-A*3004	0.006	0.003	ns	
HLA-A*3101	0.032	0.022	ns	
HLA-A*3103	0	0	-	
HLA-A*3112	0	0	-	
HLA-A*3201	0.034	0.039	ns	
HLA-A*3301	0.018	0.022	ns	
HLA-A*3303	0	0.003	ns	
HLA-A*3401	0	0	-	ns
HLA-A*3402	0.004	0.002	ns	
HLA-A*3601	0.002	0	ns	
HLA-A*3603	0	0	-	
HLA-A*6601	0.008	0.006	ns	
HLA-A*6602	0	0	-	
HLA-A*6801	0.013	0.020	ns	
HLA-A*6802	0.0108	0.013	ns	
HLA-A*6803	0	0	-	
HLA-A*6901	0.002	0.006	ns	
HLA-A*7401	0.002	0.001	ns	ns
HLA-A*8001	0.002	0.002	ns	
HLA-B*0702	0.063	0.076	ns	
HLA-B*0704	0	0	-	
HLA-B*0705	0.002	0.007	ns	ns
HLA-B*0801	0.045	0.063	ns	

HLA-B*1301	0	0.001	ns	
HLA-B*1302	0.009	0.015	ns	
HLA-B*1401	0.025	0.021	ns	
HLA-B*1402	0.041	0.043	ns	
HLA-B*1501	0.032	0.031	ns	
HLA-B*1502	0	0	-	
HLA-B*1503	0.013	0.004	0.019	3.03 (1.20-7.65)
HLA-B*1504	0	0	-	
HLA-B*1505	0	0	-	
HLA-B*1506	0	0	-	
HLA-B*1507	0	0	-	
HLA-B*1508	0	0	-	
HLA-B*1509	0	0	-	
HLA-B*1510	0.0005	0.002	ns	
HLA-B*1513	0	0	-	
HLA-B*1515	0	0	-	
HLA-B*1516	0	0.002	ns	
HLA-B*1517	0.002	0.004	ns	
HLA-B*1518	0.004	0.003	ns	
HLA-B*1521	0	0	-	
HLA-B*1524	0	0	-	
HLA-B*1525	0.005	0	ns	
HLA-B*1526	0	0	-	
HLA-B*1527	0	0	-	
HLA-B*1530	0	0	-	
HLA-B*1801	0.056	0.082	0.046	0.68 (0.46-0.99)
HLA-B*1802	0	0	-	
HLA-B*2702	0.002	0	ns	
HLA-B*2703	0	0	-	
HLA-B*2704	0	0	-	
HLA-B*2705	0.023	0.027	ns	
HLA-B*2707	0	0	-	
HLA-B*3501	0.034	0.054	0.049	0.61 (0.38-0.99)
HLA-B*3502	0.011	0.018	ns	
HLA-B*3503	0.009	0.020	ns	
HLA-B*3505	0	0	-	
HLA-B*3508	0.004	0.005	ns	
HLA-B*3512	0	0	-	
HLA-B*3514	0	0	-	
HLA-B*3517	0	0.001	ns	
HLA-B*3521	0	0	-	
HLA-B*3543	0	0	-	
HLA-B*3701	0.004	0.007	ns	
HLA-B*3801	0.023	0.036	ns	
HLA-B*3802	0	0	-	
HLA-B*3901	0.002	0.010	ns	
HLA-B*3902	0	0	-	
HLA-B*3906	0.009	0.006	ns	
HLA-B*3908	0	0	-	
HLA-B*3909	0	0	-	
HLA-B*3910	0.002	0	-	
HLA-B*4001	0.009	0.022	ns	
HLA-B*4002	0.007	0.009	ns	
HLA-B*4005	0	0	-	
HLA-B*4006	0.007	0.001	0.016	5.53 (1.38-22.26)

HLA-B*4009	0	0	-	
HLA-B*4010	0	0	-	
HLA-B*4023	0	0	-	
HLA-B*4101	0.002	0.012	ns	
HLA-B*4102	0.004	0.005	ns	
HLA-B*4201	0.002	0	-	
HLA-B*4202	0	0.001	ns	
HLA-B*4402	0.050	0.067	ns	
HLA-B*4403	0.074	0.090	ns	
HLA-B*4404	0	0	-	
HLA-B*4405	0	0.002	ns	
HLA-B*4501	0.016	0.015	ns	
HLA-B*4601	0.002	0.001	ns	
HLA-B*4701	0.002	0.001	ns	
HLA-B*4801	0	0.001	ns	
HLA-B*4901	0.040	0.040	ns	
HLA-B*5001	0.025	0.025	ns	
HLA-B*5002	0	0	-	
HLA-B*5101	0.241	0.080	4.74E-26	3.92 (3.04-5.05)
HLA-B*5102	0	0	-	
HLA-B*5105	0	0	-	
HLA-B*5106	0	0	-	
HLA-B*5108	0.004	0	ns	
HLA-B*5201	0.007	0.015	ns	
HLA-B*5301	0.014	0.011	ns	
HLA-B*5401	0	0	-	
HLA-B*5501	0.007	0.014	ns	
HLA-B*5601	0.004	0.004	ns	
HLA-B*5603	0	0	-	
HLA-B*5604	0	0	-	
HLA-B*5701	0.047	0.027	0.01	1.81 (1.14-2.86)
HLA-B*5702	0	0	-	
HLA-B*5703	0.009	0.002	0.01	4.62 (1.40-15.24)
HLA-B*5801	0.005	0.014	ns	
HLA-B*5802	0	0	-	
HLA-B*7301	0	0.001	ns	
HLA-B*7801	0	0	-	
HLA-B*8101	0	0	-	
<hr/>				
HLA-C*0102	0.025	0.034	ns	
HLA-C*0202	0.063	0.040	0.019	1.59 (1.08-2.33)
HLA-C*0206	0	0	-	
HLA-C*0302	0	0.003	ns	
HLA-C*0303	0.036	0.035	ns	
HLA-C*0304	0.014	0.029	ns	
HLA-C*0305	0	0	-	
HLA-C*0306	0	0	-	
HLA-C*0401	0.094	0.127	0.03	0.71 (0.52-0.96)
HLA-C*0403	0	0.001	ns	
HLA-C*0404	0	0	-	
HLA-C*0407	0	0	-	
HLA-C*0410	0	0	-	
HLA-C*0501	0.081	0.107	ns	
HLA-C*0602	0.083	0.065	ns	
HLA-C*0701	0.130	0.158	ns	
HLA-C*0702	0.076	0.082	ns	

HLA-C*0704	0.005	0.014	ns	
HLA-C*0726	0	0	-	
HLA-C*0801	0	0.001	ns	
HLA-C*0802	0.068	0.060	ns	
HLA-C*1202	0.007	0.014	ns	
HLA-C*1203	0.040	0.070	0.01	0.56 (0.36-0.87)
HLA-C*1204	0	0	-	
HLA-C*1402	0.049	0.018	2.34E-05	2.76 (1.72-4.42)
HLA-C*1502	0.097	0.031	3.47E-11	3.47 (2.42-4.98)
HLA-C*1504	0	0	-	
HLA-C*1505	0.005	0.009	ns	
HLA-C*1601	0.065	0.077	ns	
HLA-C*1602	0.045	0.007	2.18E-10	6.72 (3.73-12.11)
HLA-C*1604	0.009	0.001	0.002	13.9 (2.68-72.02)
HLA-C*1701	0.007	0.016	ns	
HLA-C*1801	0.001	0.001	ns	
<hr/>				
HLA-DQA1*0101	0.140	0.158	ns	
HLA-DQA1*0102	0.119	0.136	ns	
HLA-DQA1*0103	0.079	0.083	ns	
HLA-DQA1*0201	0.187	0.166	ns	
HLA-DQA1*0301	0.166	0.138	ns	
HLA-DQA1*0401	0.045	0.025	0.01	1.80 (1.14-2.83)
HLA-DQA1*0501	0.259	0.293	ns	
HLA-DQA1*0601	0.005	0.003	ns	
HLA-DQB1*0201	0.106	0.134	ns	
HLA-DQB1*0202	0.162	0.144	ns	
HLA-DQB1*0301	0.200	0.200	ns	
HLA-DQB1*0302	0.110	0.095	ns	
HLA-DQB1*0303	0.034	0.032	ns	
HLA-DQB1*0304	0.002	0.001	ns	
HLA-DQB1*0305	0	0	-	
HLA-DQB1*0401	0	0	-	
HLA-DQB1*0402	0.049	0.030	0.025	1.64 (1.06-2.52)
HLA-DQB1*0501	0.122	0.132	ns	
HLA-DQB1*0502	0.020	0.021	ns	
HLA-DQB1*0503	0.0234	0.029	ns	
HLA-DQB1*0504	0	0.001	ns	
HLA-DQB1*0601	0.011	0.013	ns	
HLA-DQB1*0602	0.061	0.071	ns	
HLA-DQB1*0603	0.068	0.074	ns	
HLA-DQB1*0604	0.027	0.028	ns	
HLA-DQB1*0609	0.005	0.009	ns	
<hr/>				
HLA-DPA1*0103	0.239	0.245	ns	
HLA-DPA1*0104	0.009	0.005	ns	
HLA-DPA1*0105	0.002	0.002	ns	
HLA-DPA1*0201	0.192	0.206	ns	
HLA-DPA1*0202	0.034	0.030	ns	
HLA-DPA1*0301	0.002	0.003	ns	
HLA-DPA1*0401	0	0	ns	
HLA-DPB1*0101	0.043	0.047	ns	
HLA-DPB1*0201	0.178	0.137	0.01	1.36 (1.07-1.74)
HLA-DPB1*0202	0.0180	0.032	ns	
HLA-DPB1*0301	0.081	0.085	ns	
HLA-DPB1*0401	0.412	0.383	ns	
HLA-DPB1*0402	0.067	0.100	0.01	0.64 (0.45-0.91)

HLA-DPB1*0501	0.022	0.017	ns	
HLA-DPB1*0601	0.009	0.015	ns	
HLA-DPB1*0901	0	0.005	ns	
HLA-DPB1*1001	0.020	0.024	ns	
HLA-DPB1*1101	0.044	0.042	ns	
HLA-DPB1*1301	0.043	0.025	0.02	1.78 (1.07-2.88)
HLA-DPB1*1401	0.007	0.031	0.003	0.22 (0.08-0.60)
HLA-DPB1*1501	0.007	0.009	ns	
HLA-DPB1*1601	0.005	0.004	ns	
HLA-DPB1*1701	0.036	0.035	ns	
HLA-DPB1*1801	0	0	-	
HLA-DPB1*2001	0	0	-	
HLA-DPB1*2101	0	0	-	
HLA-DPB1*2201	0	0	-	
HLA-DPB1*2301	0.002	0.002	ns	
HLA-DPB1*2601	0.002	0.001	ns	
HLA-DPB1*2801	0	0	-	
HLA-DPB1*3001	0	0	-	
HLA-DPB1*3101	0	0	-	
HLA-DPB1*3301	0	0	-	
HLA-DPB1*3401	0	0	-	
HLA-DPB1*3501	0	0	-	
HLA-DPB1*4001	0	0.001	ns	
HLA-DPB1*4501	0	0	-	
HLA-DPB1*4801	0	0	-	
HLA-DPB1*8101	0	0	-	
HLA-DPB1*8501	0	0	-	
<hr/>				
HLA-DRB1*0101	0.038	0.064	0.01	0.57 (0.36-0.90)
HLA-DRB1*0102	0.040	0.040	ns	
HLA-DRB1*0103	0.020	0.012	ns	
HLA-DRB1*0301	0.100	0.130	0.04	0.74 (0.55-0.99)
HLA-DRB1*0302	0.004	0.002	ns	
HLA-DRB1*0304	0	0.001	ns	
HLA-DRB1*0401	0.041	0.035	ns	
HLA-DRB1*0402	0.027	0.023	ns	
HLA-DRB1*0403	0.004	0.008	ns	
HLA-DRB1*0404	0.038	0.030	ns	
HLA-DRB1*0405	0.023	0.027	ns	
HLA-DRB1*0406	0	0.001	ns	
HLA-DRB1*0407	0.016	0.005	0.007	3.12 (1.36-7.17)
HLA-DRB1*0408	0.005	0.004	ns	
HLA-DRB1*0409	0	0.001	ns	
HLA-DRB1*0410	0	0.001	ns	
HLA-DRB1*0411	0	0.001	ns	
HLA-DRB1*0437	0	0.001	ns	
HLA-DRB1*0701	0.187	0.165	ns	
HLA-DRB1*0801	0.032	0.022	ns	
HLA-DRB1*0802	0.005	0.003	ns	
HLA-DRB1*0803	0.004	0.003	ns	
HLA-DRB1*0804	0.004	0.002	ns	
HLA-DRB1*0806	0	0.001	ns	
HLA-DRB1*0901	0.009	0.010	ns	
HLA-DRB1*1001	0.021	0.017	ns	
HLA-DRB1*1101	0.074	0.062	ns	
HLA-DRB1*1102	0.013	0.020	ns	

HLA-DRB1*1103	0.022	0.011	0.03	2.09 (1.07-4.12)
HLA-DRB1*1104	0.045	0.049	ns	
HLA-DRB1*1106	0	0.001	ns	
HLA-DRB1*1108	0	0.001	ns	
HLA-DRB1*1201	0.005	0.011	ns	
HLA-DRB1*1202	0.002	0.001	ns	0.30 (0.10-0.97)
HLA-DRB1*1301	0.068	0.069	ns	
HLA-DRB1*1302	0.032	0.038	ns	
HLA-DRB1*1303	0.005	0.018	0.04	
HLA-DRB1*1304	0	0.001	ns	
HLA-DRB1*1305	0	0	-	
HLA-DRB1*1401	0.022	0.024	ns	
HLA-DRB1*1402	0	0	-	
HLA-DRB1*1404	0.004	0.003	ns	
HLA-DRB1*1405	0	0	-	
HLA-DRB1*1406	0	0	-	
HLA-DRB1*1407	0	0	-	
HLA-DRB1*1415	0	0	-	
HLA-DRB1*1501	0.067	0.077	ns	
HLA-DRB1*1502	0.011	0.013	ns	
HLA-DRB1*1503	0	0.001	ns	
HLA-DRB1*1601	0.011	0.012	ns	
HLA-DRB1*1602	0	0.001	ns	

Suppl. Table 5. List of SNPs, amino acid and classical HLA classical allele-groups associated at GWAS level

Variation	Position in Chr 6 (GRCh36)	Allele*	P-value	OR (CI 95%)
SNP_B_31432646	31432646	A	1.64E-26	3.99(3.096-5.152)
HLA_B_51	31431272	P	2.12E-26	3.978(3.084-5.131)
HLA_B_5101	31431272	P	1.29E-25	3.892(3.017-5.02)
SNP_B_31432107	31432107	T	1.86E-22	3.354(2.629-4.277)
SNP_B_31432222	31432222	A	1.86E-22	3.354(2.629-4.277)
SNP_B_31432180_CT	31432180	A	2.32E-21	2.861(2.302-3.554)
AA_B_97_31432180_TV	31432180	P	2.32E-21	2.861(2.302-3.554)
SNP_B_31432180_C	31432180	A	2.73E-19	2.599(2.11-3.201)
AA_B_97_31432180_SRW	31432180	A	2.73E-19	2.599(2.11-3.201)
AA_B_97_31432180_T	31432180	P	4.80E-18	2.719(2.168-3.41)
SNP_B_31432180_G	31432180	P	4.80E-18	2.719(2.168-3.41)
AA_B_97_31432180_SRN	31432180	A	1.04E-17	2.341(1.927-2.844)
rs2442736	31454600	G	2.94E-17	2.439(1.983-2.999)
AA_B_97_31432180_SR	31432180	A	1.97E-16	2.222(1.837-2.688)
SNP_B_31432180_CA	31432180	A	5.91E-16	2.444(1.968-3.035)
AA_B_97_31432180_TN	31432180	P	5.91E-16	2.444(1.968-3.035)
SNP_B_31432476_A	31432476	P	2.20E-15	2.242(1.836-2.737)
AA_B_80_31432476_I	31432476	P	2.20E-15	2.242(1.836-2.737)
AA_B_97_31432180_TW	31432180	P	2.18E-14	2.176(1.782-2.656)
rs4947296	31166157	C	3.02E-13	2.617(2.021-3.389)
AA_B_97_31432180_SRV	31432180	A	3.43E-13	2.058(1.694-2.5)
rs9380217	31159532	T	3.61E-13	2.598(2.008-3.36)
rs9380215	31157634	A	3.61E-13	2.598(2.008-3.36)
rs9266395	31443545	T	4.72E-13	2.115(1.727-2.591)
rs6933050	31451611	C	4.72E-13	2.115(1.727-2.591)
rs9266381	31442771	T	5.24E-13	2.112(1.724-2.588)
rs9266399	31443785	T	5.74E-13	2.106(1.72-2.579)
rs4959053	31207556	A	1.76E-12	2.593(1.99-3.379)
HLA_C_1502	31346171	P	9.88E-12	3.501(2.441-5.021)
SNP_C_31347028	31347028	T	9.88E-12	3.501(2.441-5.021)
AA_C_116_31347029_L	31347029	P	9.88E-12	3.501(2.441-5.021)
rs9378152	31153533	G	1.55E-11	2.558(1.947-3.36)
SNP_B_31432188	31432188	A	3.62E-11	1.948(1.599-2.373)
SNP_B_31432187_A	31432187	P	3.62E-11	1.948(1.599-2.373)
SNP_B_31432186	31432186	C	3.62E-11	1.948(1.599-2.373)
AA_B_95_31432186_W	31432186	P	3.62E-11	1.948(1.599-2.373)
SNP_B_31432185	31432185	C	3.62E-11	1.948(1.599-2.373)
rs3734854	31186815	A	4.09E-11	2.341(1.819-3.014)
rs9391709	31185384	C	4.09E-11	2.341(1.819-3.014)
rs2233984	31187243	A	4.09E-11	2.341(1.819-3.014)
rs12664384	31380871	T	4.96E-11	3.06(2.192-4.272)
SNP_B_31431959	31431959	G	1.21E-10	1.901(1.563-2.311)
AA_B_171_31431958_H	31431958	P	1.21E-10	1.901(1.563-2.311)
AA_B_171_31431958_Y	31431958	A	1.21E-10	1.901(1.563-2.311)
HLA_C_1602	31346171	P	1.97E-10	6.777(3.759-12.22)
AA_B_97_31432180_STV	31432180	P	2.84E-10	1.835(1.52-2.216)
rs2516449	31512040	C	3.83E-10	1.925(1.568-2.363)
SNP_B_31432645_C	31432645	P	6.27E-10	1.824(1.508-2.207)
AA_B_24_31432644_A	31432644	P	6.27E-10	1.824(1.508-2.207)
AA_B_152_31432015_E	31432015	P	6.59E-10	1.796(1.491-2.163)
SNP_B_31432015	31432015	T	6.59E-10	1.796(1.491-2.163)
AA_B_152_31432015_V	31432015	A	6.59E-10	1.796(1.491-2.163)

rs16899166	31366075 T	6.60E-10	2.898(2.068-4.063)
rs12663108	31377534 A	8.31E-10	2.877(2.053-4.032)
AA_B_97_31432180_RW	31432180 A	8.42E-10	1.801(1.492-2.173)
SNP_C_31345897_G	31345897 A	8.48E-10	2.298(1.762-2.998)
AA_B_97_31432180_RN	31432180 A	8.83E-10	1.779(1.48-2.139)
SNP_C_31345897_A	31345897 P	1.02E-09	2.288(1.754-2.985)
AA_B_97_31432180_R	31432180 A	1.96E-09	1.76(1.463-2.117)
SNP_C_31345216	31345216 A	2.78E-09	2.79(1.989-3.913)
rs5009853	31348615 T	2.78E-09	2.79(1.989-3.913)
SNP_C_31347039	31347039 G	2.78E-09	2.79(1.989-3.913)
HLA_C_15	31346171 P	2.78E-09	2.79(1.989-3.913)
SNP_C_31346772	31346772 A	2.78E-09	2.79(1.989-3.913)
SNP_C_31346801	31346801 A	2.78E-09	2.79(1.989-3.913)
AA_C_113_31347038	31347038 H	2.78E-09	2.79(1.989-3.913)
SNP_B_31432515_A	31432515 P	1.19E-08	1.721(1.428-2.075)
AA_B_67_31432515_FM	31432515 P	1.19E-08	1.721(1.428-2.075)
rs2256028	31487177 T	4.44E-08	1.687(1.399-2.034)
rs3094584	31491827 T	4.70E-08	1.712(1.412-2.076)
SNP_A_30018709	30018709 A	6.94E-08	1.641(1.371-1.965)
AA_A_66_30018708	30018708 K	6.94E-08	1.641(1.371-1.965)
AA_A_62_30018696_GE	30018696 P	6.99E-08	1.643(1.372-1.968)
SNP_A_30018695	30018695 G	6.99E-08	1.643(1.372-1.968)
SNP_A_30018506	30018506 T	7.47E-08	1.64(1.369-1.964)
AA_B_194_31431315_I	31431315 A	7.60E-08	1.723(1.413-2.101)
AA_B_194_31431315_V	31431315 P	7.60E-08	1.723(1.413-2.101)
SNP_B_31431316	31431316 C	7.60E-08	1.723(1.413-2.101)
SNP_B_31431447	31431447 T	7.60E-08	1.723(1.413-2.101)
rs2250264	30929166 A	9.06E-08	1.719(1.409-2.097)
rs4711206	30095163 A	9.88E-08	1.715(1.406-2.091)
rs9393989	30148063 A	9.88E-08	1.715(1.406-2.091)
rs9380150	30118471 C	9.92E-08	1.704(1.4-2.073)

Suppl. Table 6 Study of replicacion of the classical HLA-A and B allele group at two-digits level

Allele	Allele Frequency				Allele Frequency				Allele Frequency			
	BD	Controls	P-value	OR (95%CI)	BD	Controls	P-value	OR (95%CI)	BD	Controls	P-value	OR (95%CI)
	n=278	n=1517			n=129	n=377			n=407	n=1894		
HLA-A*01	0.103	0.123	ns		0.105	0.118			0.098	0.118		
HLA-A*02	0.344	0.247	3.09E-06	1.60 (1.32-1.94)	0.345	0.249	0.000	1.59 (1.22-	0.349	0.249	1.31E-08	1.62 (1.38-
HLA-A*03	0.063	0.114	0.0004	0.52 (0.36-0.75)	0.062	0.113	0.014	0.52 (0.31-	0.063	0.113	5.16E-05	0.53 (0.39-
HLA-A*11	0.059	0.064	ns		0.050	0.066	ns		0.058	0.066	ns	
HLA-A*23	0.038	0.028	ns		0.043	0.030	ns		0.038	0.030	ns	
HLA-A*24	0.11	0.093	ns		0.097	0.090	ns		0.104	0.090	ns	
HLA-A*25	0.017	0.001	0.0000243		0.004	0.003	ns				ns	
HLA-A*26	0.034	0.063	0.009		0.047	0.059	ns				ns	
HLA-A*29	0.072	0.08	ns		0.050	0.080	ns		0.067	0.080	ns	
HLA-A*30	0.034	0.056	0.037	0.60 (0.37-0.97)	0.035	0.055	ns		0.037	0.055	0.04	0.66 (0.45-
HLA-A*31	0.032	0.022	ns		0.039	0.023	ns		0.033	0.023	ns	
HLA-A*32	0.034	0.039	ns		0.031	0.039	ns		0.033	0.039	ns	
HLA-A*33	0.018	0.026	ns		0.016	0.028	ns		0.018	0.028	ns	
HLA-A*34	0.004	0.002	ns		0	0.002	-		0.003	0.002	ns	
HLA-A*36	0.002	0	ns		0	0.001	-		0.001	0.001	ns	
HLA-A*66	0.009	0.006	ns		0.012	0.001	ns		0.010	0.006	ns	
HLA-A*68	0.002	0.032	ns		0.066	0.037	0.02	1.85 (1.10-	0.036	0.036	ns	
HLA-A*69	0.002	0.006	ns		0	0.006	-		0.001	0.006	ns	
HLA-A*74	0.002	0.001	ns		0	0.001	-		0	0.001	-	
HLA-A*80	0.002	0.002	ns		0	0.002	-		0.001	0.002	ns	
HLA-B*07	0.065	0.082	ns		0.063	0.079	ns		0.067	0.079	ns	
HLA-B*08	0.045	0.063	ns		0.046	0.059	ns		0.042	0.059	ns	
HLA-B*13	0.009	0.016	ns		0.021	0.015	ns		0.012	0.015	ns	
HLA-B*14	0.067	0.064	ns		0.054	0.069	ns		0.063	0.069	ns	
HLA-B*15	0.061	0.047	ns		0.083	0.045	0.009	1.91 (1.18-	0.063	0.045	0.043	1.41 (1.02-1.96)
HLA-B*18	0.058	0.082	0.046	0.68 (0.47-0.99)	0.05	0.082	ns		0.055	0.082	0.010	0.65 (0.47-0.91)
HLA-B*27	0.025	0.027	ns		0.042	0.028	ns		0.032	0.027	ns	
HLA-B*35	0.058	0.099	0.002	0.55 (0.38-0.81)	0.075	0.103	ns		0.065	0.103	0.001	0.61 (0.45-0.82)
HLA-B*37	0.004	0.007	ns		0.013	0.007	ns		0.005	0.007	ns	
HLA-B*38	0.023	0.036	ns		0.021	0.037	ns		0.022	0.037	0.035	0.58 (0.35-0.96)
HLA-B*39	0.013	0.016	ns		0.004	0.016	ns		0.010	0.016	ns	
HLA-B*40	0.023	0.033	ns		0.021	0.033	ns		0.022	0.033	ns	
HLA-B*41	0.005	0.017	ns		0.125	0.016	ns		0.008	0.016	ns	
HLA-B*42	0.002	0.001	ns		0	0.001	-		0.001	0.001	ns	
HLA-B*44	0.124	0.16	0.031	0.74 (0.57-0.98)	0.133	0.157	ns		0.132	0.157	ns	
HLA-B*45	0.016	0.015	ns		0.008	0.017	ns		0.014	0.017	ns	
HLA-B*46	0.002	0.001	ns		0.004	0.001	ns		0.001	0.001	ns	
HLA-B*47	0.002	0.001	ns		0	0.001	-		0.001	0.001	ns	
HLA-B*48	0	0.001	-		0	0.001	-		0	0.001	-	
HLA-B*49	0.036	0.04	ns		0.013	0.042	0.039	0.29 (0.09-	0.03	0.04	ns	
HLA-B*50	0.025	0.025	ns		0.0123	0.026	ns		0.021	0.026	ns	
HLA-B*51	0.245	0.08	7.9E-27	3.72 (2.94-4.70)	0.217	0.079	1.37E-	3.23 (2.32-	0.231	0.079	1.01E-31	3.50 (2.85-4.29)
HLA-B*52	0.007	0.015	ns		0.008	0.015	ns		0.008	0.015	ns	
HLA-B*53	0.014	0.011	ns		0.021	0.013	ns		0.017	0.013	ns	
HLA-B*54	0	0	-		0	0	-		0	0	-	
HLA-B*55	0.007	0.014	ns		0.021	0.012	ns		0.010	0.012	ns	
HLA-B*56	0.004	0.004	ns		0.004	0.004	ns		0.005	0.004	ns	
HLA-B*57	0.058	0.028	0.001	2.00 (1.31-3.05)	0.04	0.027	ns		0.051	0.027	0.0007	1.93 (1.33-2.818]
HLA-B*58	0.005	0.014	ns		0	0.014	-		0.005	0.015	0.046	0.36 (0.13-0.99)
HLA-B*73	0	0.001	-		0	0.001	-		0	0.001	ns	
HLA-B*78	0	0	-		0	0	-		0	0	ns	
HLA-B*81	0	0	-		0	0	-		0	0	ns	

Suppl. Table 7 Multivariate analysis including the classical HLA-A and B allele-groups associated in the overall cohort

HLA	P-Value	OR	95% CI
A*02	0.0643	1.25	0.99-1.91
A*03	0.0009	0.55	0.39-0.78
A*30	0.3867	0.82	0.52-1.29
B*15	$<10^{-4}$	2.23	1.56-3.17
B*18	0.362	0.83	0.57-1.21
B*35	0.1416	0.77	0.55-1.09
B*38	0.4795	0.82	0.48-1.41
B*51	$<10^{-4}$	4.30	3.33-5.56
B*57	0.0001	2..30	1.51-3.51
B*58	0..525	0.47	0.16-1.32

Suppl. Table 7. Allelic frequencies of HLA*02 and A*03 in different group of patients and controls

	HLA-B*51 NEGATIVE			
	BD N=225	CONTROL S N=1607	P	OR
A*02	137 (0.28)	738 (0.23)	0.0004	1.47
A*03	36 (0.08)	360 (0.11)	0.04	0.69
	HLA-B*51 POSITIVE			
	N=168	N=282	P	OR
A*02	132 (0.39)	201 (0.36)	>0.05	
A*03	14 (0.04)	62 (0.11)	0.0004	0.35
	HLA-B*15+51+57 NEGATIVE			
	N=157	N=1356		
A*02	97 (0.31)	630 (0.23)	0.003	1.48
A*03	29 (0.09)	314 (0.12)	>0.05	
	HLA-B*15 or 51 or 57 POSITIVE			
	N=193	N=380		
A*02	143 (0.37)	241 (0.32)	>0.05	
A*03	18 (0.05)	72 (0.09)	0.004	0.47

Suppl. Table 7. Omnibus analysis of the amino acid positions in the different HLA loci

HLA molecule	Aminoacid position	Center codon position (GRCh36)	Tested Alleles	P-value	P-value (conditioned by B 97)	P-value (conditioned by B 97 and A 66)
B	97	31432180	6	9.90E-17	NA	NA
B	80	31432476	3	1.43E-14	2.76E-03	3.85E-03
B	171	31431958	2	1.90E-10	6.35E-03	1.41E-02
B	95	31432186	3	3.51E-10	2.75E-01	4.16E-01
C	116	31347029	4	7.11E-10	2.90E-03	5.70E-03
B	152	31432015	2	7.21E-09	1.10E-03	5.33E-03
B	81	31432473	2	5.99E-08	6.34E-02	5.82E-02
B	24	31432644	3	7.87E-08	7.84E-01	6.99E-01
B	77	31432485	3	3.06E-07	4.62E-02	4.12E-02
C	113	31347038	2	4.98E-07	6.55E-02	1.02E-01
B	194	31431315	2	5.08E-07	2.06E-01	1.91E-01
B	82	31432470	2	8.91E-07	1.25E-01	1.14E-01
B	83	31432467	2	8.91E-07	1.25E-01	1.14E-01
A	66	30018708	2	1.70E-06	1.67E-05	NA
B	67	31432515	5	2.19E-06	9.21E-02	7.37E-02
A	116	30019099	3	3.13E-06	1.18E-04	1.30E-01
A	95	30019036	3	5.10E-06	1.58E-04	4.86E-02
A	114	30019093	4	7.82E-06	2.86E-04	1.86E-01
B	163	31431982	3	8.44E-06	1.47E-03	2.55E-03
C	73	31347408	2	2.79E-05	7.83E-03	1.21E-02
A	62	30018696	5	3.40E-05	2.03E-04	4.86E-01
A	74	30018732	2	3.49E-05	9.11E-04	8.18E-01
A	127	30019132	2	4.51E-05	7.12E-04	9.62E-02
A	-15	30018338	2	5.35E-05	9.28E-04	1.27E-01
A	97	30019042	3	7.89E-05	8.12E-04	1.88E-01
A	276	30020259	2	9.61E-05	2.94E-04	1.19E-01
A	321	30020836	2	9.61E-05	2.94E-04	1.19E-01
A	107	30019072	2	1.10E-04	3.73E-03	7.71E-01
B	45	31432581	4	2.12E-04	4.85E-01	5.50E-01
A	152	30019207	5	2.27E-04	1.69E-03	2.34E-01
B	-16	31432889	2	3.48E-04	1.01E-02	2.02E-02
B	116	31432123	5	3.51E-04	4.26E-02	7.28E-02
B	103	31432162	2	4.02E-04	2.07E-03	3.70E-03
A	70	30018720	2	4.65E-04	1.08E-03	3.16E-01
B	305	31430889	2	4.87E-04	5.83E-02	5.71E-02
B	177	31431940	2	7.87E-04	1.95E-01	1.34E-01
B	180	31431931	2	7.87E-04	1.95E-01	1.34E-01
C	21	31347564	2	8.75E-04	2.66E-01	3.93E-01
C	1	31347624	2	9.03E-04	6.78E-02	9.44E-02
A	161	30019234	2	1.10E-03	6.92E-03	1.16E-01
A	142	30019177	2	1.17E-03	2.42E-02	5.23E-01
A	145	30019186	2	1.17E-03	2.42E-02	5.23E-01
B	131	31432078	2	1.49E-03	2.57E-01	2.34E-01

B	-10	31432871	2	1.49E-03	6.52E-01	4.99E-01
B	114	31432129	3	2.51E-03	9.23E-01	8.11E-01
B	-8	31432865	2	2.55E-03	6.45E-01	5.51E-01
B	11	31432683	2	2.88E-03	5.79E-01	5.23E-01
B	282	31430958	2	3.03E-03	2.11E-01	2.47E-01
B	325	31430282	2	3.35E-03	1.54E-01	1.88E-01
C	-17	31347805	2	3.95E-03	1.30E-01	1.20E-01
B	12	31432680	2	3.96E-03	4.86E-01	3.91E-01
A	207	30019950	2	4.79E-03	1.95E-02	1.32E-01
A	193	30019908	2	5.25E-03	2.18E-02	1.43E-01
A	194	30019911	2	5.25E-03	2.18E-02	1.43E-01
C	114	31347035	2	5.35E-03	5.97E-01	5.28E-01
C	77	31347396	2	5.45E-03	4.27E-02	1.12E-01
C	80	31347387	2	5.45E-03	4.27E-02	1.12E-01
DPB1	178	33160960	2	5.81E-03	9.76E-03	6.48E-03
B	-11	31432874	2	7.57E-03	1.65E-01	2.81E-01
A	294	30020313	2	7.83E-03	3.12E-02	1.55E-01
B	41	31432593	2	8.17E-03	5.99E-01	5.43E-01
C	275	31345841	3	8.72E-03	6.43E-01	6.95E-01
C	253	31346031	2	8.94E-03	3.25E-01	2.54E-01
C	267	31345989	2	8.94E-03	3.25E-01	2.54E-01
B	9	31432689	3	9.28E-03	9.99E-01	9.23E-01
A	184	30019881	2	9.55E-03	3.43E-02	2.46E-01
B	74	31432494	2	1.00E-02	1.12E-03	7.05E-04
A	253	30020088	2	1.04E-02	3.44E-02	1.98E-01
C	90	31347357	2	1.07E-02	2.84E-01	2.53E-01
C	193	31346211	2	1.15E-02	3.59E-03	7.18E-04
B	-23	31432910	2	1.16E-02	3.90E-01	3.25E-01
B	-21	31432904	2	1.16E-02	3.90E-01	3.25E-01
A	17	30018561	2	1.22E-02	5.49E-02	2.86E-01
B	156	31432003	4	1.26E-02	6.90E-01	5.49E-01
C	306	31345748	2	1.30E-02	4.63E-01	3.41E-01
B	32	31432620	2	1.35E-02	5.80E-01	7.25E-01
B	178	31431937	2	1.42E-02	5.18E-01	3.24E-01
DQA1	175	32717988	3	1.43E-02	1.81E-01	1.14E-01
C	94	31347095	2	1.48E-02	8.66E-01	9.37E-01
DPB1	57	33156583	2	1.61E-02	6.04E-03	1.35E-02
C	147	31346936	2	1.64E-02	5.70E-01	4.40E-01
C	303	31345757	2	1.66E-02	3.98E-01	4.59E-01
B	94	31432189	2	1.76E-02	1.94E-01	3.11E-01
DRB1	37	32660037	5	1.77E-02	9.29E-02	1.16E-01
C	211	31346157	2	1.95E-02	6.39E-02	1.06E-01
DPB1	9	33156439	3	1.97E-02	1.11E-02	2.07E-02
C	326	31345141	2	2.00E-02	2.88E-01	2.40E-01
C	339	31345102	2	2.00E-02	2.88E-01	2.40E-01
DPB1	76	33156640	3	2.01E-02	1.08E-02	2.72E-02
C	14	31347585	2	2.12E-02	4.76E-01	5.97E-01

C	49	31347480	2	2.12E-02	4.76E-01	5.97E-01
DRB1	98	32657584	2	2.21E-02	1.28E-01	5.53E-02
DRB1	104	32657566	2	2.21E-02	1.28E-01	5.53E-02
C	95	31347092	3	2.23E-02	4.45E-01	5.63E-01
A	99	30019048	2	2.30E-02	2.01E-02	1.49E-01
C	285	31345811	3	2.33E-02	6.62E-01	5.42E-01
C	261	31346007	2	2.35E-02	3.84E-01	3.24E-01
C	273	31345971	2	2.35E-02	3.84E-01	3.24E-01
A	76	30018738	3	2.39E-02	3.47E-02	2.62E-01
DQA1	69	32717257	3	2.58E-02	2.30E-01	2.07E-01
DPB1	65	33156607	2	2.74E-02	1.30E-02	2.73E-02
B	62	31432530	2	2.78E-02	7.33E-01	6.84E-01
C	184	31346238	3	2.84E-02	7.13E-01	6.04E-01
C	97	31347086	2	3.05E-02	4.28E-01	2.78E-01
DQA1	215	32718464	2	3.08E-02	1.53E-01	7.34E-02
B	143	31432042	2	3.30E-02	2.00E-01	2.07E-01
B	147	31432030	2	3.30E-02	2.00E-01	2.07E-01
DQB1	140	32737868	2	3.35E-02	5.00E-01	7.14E-01
DQB1	182	32737742	2	3.35E-02	5.00E-01	7.14E-01
C	305	31345751	2	3.45E-02	5.64E-01	4.46E-01
C	307	31345745	2	3.45E-02	5.64E-01	4.46E-01
DPB1	36	33156520	2	3.51E-02	2.12E-02	3.42E-02
DQB1	221	32737116	2	3.84E-02	6.44E-02	6.25E-02
C	295	31345781	2	3.99E-02	5.80E-01	4.66E-01
C	16	31347579	2	4.17E-02	1.04E-01	1.86E-01
DRB1	120	32657518	2	4.24E-02	8.41E-02	7.99E-02
DQB1	185	32737733	2	4.25E-02	3.67E-01	3.76E-01
DQA1	11	32717083	2	4.32E-02	1.35E-01	9.30E-02
C	194	31346208	2	4.33E-02	7.16E-01	5.88E-01
DQA1	50	32717200	3	4.59E-02	2.19E-01	1.09E-01
DQA1	53	32717209	3	4.59E-02	2.19E-01	1.09E-01
A	63	30018699	3	4.65E-02	4.90E-02	9.91E-02
B	199	31431300	2	4.81E-02	8.98E-01	9.06E-01
DQB1	220	32737119	2	4.87E-02	8.44E-02	8.02E-02
C	-9	31347781	2	5.01E-02	7.42E-01	6.19E-01
DQB1	203	32737170	2	5.19E-02	1.51E-01	1.38E-01
B	63	31432527	2	5.70E-02	3.38E-01	3.84E-01
C	219	31346133	2	5.77E-02	3.64E-02	4.61E-02
A	79	30018747	2	6.01E-02	2.48E-02	3.91E-01
A	80	30018750	2	6.01E-02	2.48E-02	3.91E-01
A	81	30018753	2	6.01E-02	2.48E-02	3.91E-01
A	82	30018756	2	6.01E-02	2.48E-02	3.91E-01
A	83	30018759	2	6.01E-02	2.48E-02	3.91E-01
C	270	31345980	2	6.18E-02	3.04E-01	2.73E-01
C	66	31347429	2	6.56E-02	3.82E-01	4.51E-01
B	70	31432506	4	6.56E-02	7.69E-02	4.92E-02
DRB1	133	32657479	2	6.59E-02	1.78E-02	2.13E-02

DRB1	142	32657452	2	6.59E-02	1.78E-02	2.13E-02
C	156	31346909	5	6.71E-02	1.33E-01	6.74E-02
DQA1	18	32717104	2	6.72E-02	1.82E-01	1.47E-01
DQA1	45	32717185	2	6.72E-02	1.82E-01	1.47E-01
DQA1	48	32717194	2	6.72E-02	1.82E-01	1.47E-01
DQA1	55	32717215	2	6.72E-02	1.82E-01	1.47E-01
DQA1	61	32717233	2	6.72E-02	1.82E-01	1.47E-01
DQA1	64	32717242	2	6.72E-02	1.82E-01	1.47E-01
DQA1	66	32717248	2	6.72E-02	1.82E-01	1.47E-01
DQA1	56	32717218	3	6.81E-02	2.31E-01	1.74E-01
DQA1	76	32717278	3	6.81E-02	2.31E-01	1.74E-01
B	46	31432578	2	7.17E-02	6.32E-01	7.93E-01
DQA1	80	32717290	2	7.21E-02	1.89E-01	1.57E-01
DQA1	218	32718473	2	7.21E-02	1.89E-01	1.57E-01
DRB1	11	32660115	5	7.40E-02	7.63E-02	6.84E-02
A	77	30018741	3	7.50E-02	1.94E-01	6.79E-02
DRB1	-17	32665460	3	7.64E-02	4.45E-02	4.03E-02
DPB1	56	33156580	2	7.76E-02	3.53E-02	4.74E-02
DQB1	53	32740678	2	8.07E-02	2.12E-01	1.75E-01
A	73	30018729	2	8.09E-02	1.37E-01	6.02E-02
C	284	31345814	2	8.11E-02	4.00E-01	3.46E-01
C	289	31345799	2	8.11E-02	4.00E-01	3.46E-01
C	291	31345793	2	8.11E-02	4.00E-01	3.46E-01
C	308	31345742	2	8.11E-02	4.00E-01	3.46E-01
C	309	31345739	2	8.11E-02	4.00E-01	3.46E-01
DQB1	84	32740585	2	8.12E-02	2.10E-01	1.77E-01
DQB1	85	32740582	2	8.12E-02	2.10E-01	1.77E-01
DQB1	89	32740570	2	8.12E-02	2.10E-01	1.77E-01
DQB1	90	32740567	2	8.12E-02	2.10E-01	1.77E-01
DQB1	70	32740627	3	8.15E-02	4.25E-01	4.29E-01
B	167	31431970	2	8.21E-02	7.55E-01	7.96E-01
DRB1	180	32657338	3	8.26E-02	4.20E-01	3.99E-01
DPB1	96	33160714	2	8.33E-02	4.93E-02	6.74E-02
DPB1	170	33160936	2	8.33E-02	4.93E-02	6.74E-02
B	30	31432626	2	8.69E-02	8.54E-01	8.37E-01
DQA1	129	32717850	2	8.73E-02	4.49E-01	4.72E-01
B	66	31432518	2	8.75E-02	3.75E-01	3.80E-01
B	65	31432521	2	9.02E-02	3.92E-01	3.85E-01
DQA1	187	32718380	2	9.38E-02	2.21E-01	2.13E-01
DQA1	47	32717191	4	9.61E-02	3.62E-01	2.17E-01
DRB1	26	32660070	3	9.64E-02	6.52E-01	7.32E-01
DRB1	33	32660049	2	9.74E-02	1.92E-01	1.77E-01
A	282	30020277	2	9.79E-02	5.94E-02	3.13E-01
A	311	30020364	2	9.79E-02	5.94E-02	3.13E-01
DRB1	77	32659917	2	9.96E-02	7.16E-01	8.56E-01
DRB1	181	32657335	3	1.01E-01	6.17E-01	4.26E-01
DRB1	13	32660109	6	1.02E-01	2.15E-01	1.82E-01

DQB1	56	32740669	2	1.02E-01	3.97E-01	5.50E-01
C	103	31347068	2	1.03E-01	1.06E-01	1.26E-01
C	173	31346858	2	1.03E-01	1.06E-01	1.26E-01
DQA1	26	32717128	2	1.04E-01	2.33E-01	2.30E-01
A	65	30018705	2	1.07E-01	5.79E-02	9.32E-01
B	113	31432132	2	1.08E-01	3.78E-01	5.26E-01
B	158	31431997	2	1.10E-01	4.44E-01	4.42E-01
DQA1	52	32717206	3	1.12E-01	3.21E-01	1.98E-01
DPB1	55	33156577	3	1.13E-01	7.76E-02	1.16E-01
A	144	30019183	2	1.14E-01	2.87E-01	5.05E-02
C	-18	31347808	2	1.14E-01	5.03E-01	4.56E-01
C	-5	31347769	2	1.14E-01	5.03E-01	4.56E-01
C	143	31346948	2	1.14E-01	5.03E-01	4.56E-01
C	170	31346867	2	1.14E-01	5.03E-01	4.56E-01
C	248	31346046	2	1.22E-01	7.05E-03	1.55E-02
DRB1	71	32659935	4	1.22E-01	1.44E-01	2.09E-01
C	24	31347555	2	1.23E-01	9.93E-02	8.75E-02
DRB1	70	32659938	3	1.24E-01	2.44E-01	3.25E-01
DRB1	47	32660007	2	1.27E-01	3.49E-01	2.03E-01
A	245	30020064	2	1.36E-01	1.12E-01	3.55E-01
DPB1	11	33156445	2	1.36E-01	5.98E-02	1.26E-01
DRB1	96	32657590	4	1.40E-01	1.93E-01	1.87E-01
C	177	31346846	2	1.41E-01	9.07E-01	7.43E-01
DRB1	74	32659926	5	1.44E-01	9.14E-01	7.77E-01
C	163	31346888	3	1.47E-01	1.46E-01	2.21E-01
DQA1	34	32717152	2	1.50E-01	2.30E-01	1.11E-01
A	283	30020280	2	1.54E-01	8.68E-02	2.58E-01
C	-15	31347799	2	1.57E-01	1.25E-01	2.69E-01
DPB1	8	33156436	2	1.59E-01	9.66E-02	1.93E-01
DRB1	4	32665400	2	1.64E-01	2.91E-01	1.84E-01
DQA1	207	32718440	2	1.65E-01	4.47E-01	4.37E-01
A	105	30019066	2	1.66E-01	1.07E-01	3.55E-01
A	163	30019240	2	1.73E-01	1.05E-01	5.62E-01
DRB1	86	32659890	2	1.75E-01	6.01E-01	4.89E-01
A	-11	30018350	2	1.80E-01	2.92E-01	1.73E-02
DQB1	125	32737913	3	1.81E-01	3.75E-01	2.91E-01
DRB1	32	32660052	2	1.86E-01	9.75E-01	9.94E-01
DQB1	87	32740576	3	1.99E-01	3.31E-01	2.24E-01
DQB1	86	32740579	3	2.05E-01	3.75E-01	3.24E-01
DQB1	224	32737107	2	2.07E-01	1.82E-01	2.18E-01
DRB1	16	32660100	2	2.07E-01	9.17E-01	9.65E-01
DQB1	-21	32742328	3	2.09E-01	3.85E-01	2.83E-01
DQB1	-6	32742283	3	2.09E-01	3.85E-01	2.83E-01
DQB1	-4	32742277	3	2.09E-01	3.85E-01	2.83E-01
DRB1	67	32659947	3	2.13E-01	6.90E-01	8.62E-01
DRB1	14	32660106	2	2.16E-01	4.66E-01	2.46E-01
DRB1	25	32660073	2	2.16E-01	4.66E-01	2.46E-01

B	145	31432036	2	2.18E-01	1.79E-03	4.88E-03
DRB1	78	32659914	2	2.24E-01	5.53E-01	3.14E-01
DQA1	75	32717275	2	2.25E-01	5.82E-01	4.73E-01
C	304	31345754	2	2.29E-01	5.12E-01	4.37E-01
B	109	31432144	2	2.31E-01	5.35E-01	3.59E-01
DRB1	58	32659974	2	2.31E-01	3.40E-01	5.96E-01
DPB1	84	33156664	2	2.32E-01	1.62E-01	2.44E-01
C	6	31347609	2	2.35E-01	9.83E-03	2.23E-02
B	99	31432174	2	2.35E-01	4.21E-01	4.34E-01
DQA1	107	32717784	2	2.40E-01	6.00E-01	4.97E-01
DQA1	156	32717931	2	2.40E-01	6.00E-01	4.97E-01
DQA1	161	32717946	2	2.40E-01	6.00E-01	4.97E-01
DQA1	163	32717952	2	2.40E-01	6.00E-01	4.97E-01
DRB1	166	32657380	2	2.49E-01	1.90E-01	2.25E-01
A	186	30019887	2	2.51E-01	2.99E-01	2.59E-01
DRB1	-24	32665481	3	2.52E-01	3.32E-01	2.57E-01
DPB1	85	33156667	2	2.68E-01	2.04E-01	3.87E-01
DPB1	86	33156670	2	2.68E-01	2.04E-01	3.87E-01
DPB1	87	33156673	2	2.68E-01	2.04E-01	3.87E-01
DQA1	54	32717212	2	2.69E-01	5.46E-01	3.05E-01
DQA1	-16	32713236	2	2.71E-01	3.31E-01	2.57E-01
A	56	30018678	2	2.72E-01	4.72E-01	6.79E-01
A	297	30020322	2	2.78E-01	3.64E-01	3.34E-01
DRB1	140	32657458	2	2.81E-01	1.52E-01	2.61E-01
A	-22	30018317	2	2.90E-01	2.90E-01	2.04E-01
C	9	31347600	4	2.91E-01	5.84E-02	9.39E-02
DQB1	71	32740624	4	2.93E-01	6.25E-01	5.36E-01
DQB1	55	32740672	3	2.96E-01	5.51E-01	3.68E-01
A	90	30018780	2	3.06E-01	1.88E-01	3.55E-01
B	275	31430979	2	3.11E-01	2.42E-01	1.89E-01
B	295	31430919	2	3.11E-01	2.42E-01	1.89E-01
B	296	31430916	2	3.11E-01	2.42E-01	1.89E-01
B	298	31430910	2	3.11E-01	2.42E-01	1.89E-01
B	299	31430907	2	3.11E-01	2.42E-01	1.89E-01
B	300	31430904	2	3.11E-01	2.42E-01	1.89E-01
A	334	30021017	2	3.14E-01	4.87E-01	2.30E-02
DQB1	135	32737883	2	3.25E-01	2.54E-01	1.27E-01
A	299	30020328	2	3.25E-01	4.92E-01	1.92E-02
DQB1	77	32740606	2	3.29E-01	7.64E-01	4.14E-01
DQB1	38	32740723	2	3.30E-01	9.06E-01	6.38E-01
C	35	31347522	2	3.30E-01	7.63E-01	9.89E-01
A	156	30019219	4	3.33E-01	3.00E-01	7.38E-01
DQB1	-5	32742280	4	3.40E-01	5.63E-01	3.75E-01
B	69	31432509	2	3.41E-01	1.50E-02	9.30E-03
DRB1	31	32660055	3	3.53E-01	3.94E-01	4.83E-01
B	71	31432503	2	3.57E-01	1.71E-02	1.02E-02
A	9	30018537	4	3.58E-01	3.38E-01	1.19E-01

DQB1	116	32737940	2	3.60E-01	5.87E-01	7.06E-01
C	138	31346963	2	3.61E-01	7.19E-01	9.54E-01
DPB1	69	33156619	3	3.68E-01	3.74E-01	3.33E-01
DRB1	9	32660121	2	3.69E-01	1.63E-01	2.97E-01
DRB1	-1	32665412	3	3.69E-01	7.56E-02	6.71E-02
A	298	30020325	2	3.70E-01	4.93E-01	4.21E-02
A	307	30020352	2	3.70E-01	4.93E-01	4.21E-02
B	239	31431180	2	3.78E-01	2.70E-01	2.29E-01
B	253	31431138	2	3.78E-01	2.70E-01	2.29E-01
B	267	31431096	2	3.78E-01	2.70E-01	2.29E-01
B	268	31431093	2	3.78E-01	2.70E-01	2.29E-01
B	270	31431087	2	3.78E-01	2.70E-01	2.29E-01
A	288	30020295	2	3.81E-01	4.53E-01	2.97E-01
DQB1	30	32740747	3	3.82E-01	6.41E-01	4.24E-01
DRB1	40	32660028	2	3.96E-01	2.73E-01	3.64E-01
DQB1	14	32740795	2	3.97E-01	6.46E-01	7.56E-01
A	151	30019204	2	4.02E-01	9.84E-01	1.38E-01
C	11	31347594	2	4.07E-01	2.54E-01	2.92E-01
DRB1	30	32660058	5	4.08E-01	4.40E-01	4.49E-01
DQA1	25	32717125	2	4.12E-01	9.09E-01	9.42E-01
C	152	31346921	2	4.16E-01	7.09E-02	5.51E-02
DPB1	35	33156517	3	4.20E-01	4.06E-01	3.91E-01
DRB1	38	32660034	3	4.21E-01	2.35E-01	3.43E-01
DRB1	149	32657431	2	4.39E-01	6.50E-01	5.29E-01
B	245	31431162	2	4.42E-01	5.33E-01	5.89E-01
A	246	30020067	2	4.73E-01	7.32E-01	3.43E-02
DPA1	111	33144977	2	5.04E-01	5.04E-01	6.00E-01
DPA1	127	33144929	2	5.04E-01	5.04E-01	6.00E-01
DPA1	160	33144830	2	5.04E-01	5.04E-01	6.00E-01
DPA1	228	33144412	2	5.04E-01	5.04E-01	6.00E-01
DPA1	83	33145401	2	5.04E-01	5.07E-01	6.32E-01
A	166	30019249	2	5.05E-01	4.62E-01	9.93E-01
A	167	30019252	2	5.05E-01	4.62E-01	9.93E-01
DPA1	31	33145557	2	5.06E-01	5.04E-01	6.49E-01
DPA1	50	33145500	2	5.06E-01	5.04E-01	6.49E-01
C	99	31347080	4	5.11E-01	7.54E-02	1.52E-01
DQB1	167	32737787	2	5.26E-01	5.25E-01	3.82E-01
A	150	30019201	2	5.31E-01	3.60E-01	7.58E-01
A	158	30019225	2	5.31E-01	3.60E-01	7.58E-01
DQB1	45	32740702	2	5.35E-01	9.19E-01	6.41E-01
A	44	30018642	2	5.39E-01	3.65E-01	7.51E-01
A	67	30018711	2	5.39E-01	3.65E-01	7.51E-01
DRB1	112	32657542	2	5.59E-01	6.57E-01	5.51E-01
DRB1	60	32659968	3	5.59E-01	9.11E-01	6.23E-01
DQB1	13	32740798	2	5.67E-01	5.27E-01	3.69E-01
DQB1	130	32737898	2	5.73E-01	7.86E-01	7.98E-01
DRB1	10	32660118	3	5.93E-01	4.78E-01	5.92E-01

A	-20	30018323	2	6.08E-01	7.39E-01	5.98E-01
A	310	30020361	2	6.08E-01	7.39E-01	5.98E-01
A	314	30020815	2	6.08E-01	7.39E-01	5.98E-01
DQB1	3	32742259	2	6.11E-01	6.12E-02	1.15E-01
DQB1	126	32737910	2	6.11E-01	5.73E-01	7.20E-01
DRB1	12	32660112	2	6.12E-01	9.22E-01	7.83E-01
DPB1	215	33161618	2	6.18E-01	6.73E-01	5.11E-01
A	149	30019198	2	6.20E-01	7.07E-01	6.09E-01
DPA1	11	33145617	2	6.49E-01	7.84E-01	8.43E-01
DQA1	41	32717173	2	6.67E-01	2.69E-01	1.74E-01
DQB1	26	32740759	3	6.77E-01	6.91E-01	5.79E-01
DRB1	28	32660064	2	6.80E-01	5.41E-01	3.55E-01
DRB1	85	32659893	2	6.80E-01	4.01E-01	4.52E-01
A	43	30018639	2	6.82E-01	5.97E-01	8.90E-01
DRB1	231	32656010	3	6.91E-01	3.99E-01	4.26E-01
DQB1	-27	32742346	3	6.98E-01	5.84E-01	5.67E-01
DQB1	-9	32742292	3	6.98E-01	5.84E-01	5.67E-01
A	-2	30018377	2	6.99E-01	7.53E-01	5.91E-01
A	255	30020094	2	7.07E-01	6.05E-01	4.73E-01
A	268	30020133	2	7.07E-01	6.05E-01	4.73E-01
DQA1	130	32717853	2	7.12E-01	2.85E-01	1.93E-01
DQB1	57	32740666	4	7.25E-01	3.39E-01	1.51E-01
DQA1	40	32717170	2	7.39E-01	9.22E-01	7.17E-01
DQA1	51	32717203	2	7.39E-01	9.22E-01	7.17E-01
DQB1	197	32737188	2	7.44E-01	6.00E-02	1.17E-01
A	109	30019078	2	7.46E-01	8.30E-01	6.49E-01
DQB1	37	32740726	3	7.47E-01	1.28E-01	1.51E-01
DQB1	74	32740615	3	7.48E-01	6.36E-01	4.09E-01
DQB1	75	32740612	2	7.50E-01	5.08E-01	2.83E-01
DRB1	57	32659977	4	7.53E-01	8.93E-01	6.82E-01
DQB1	66	32740639	2	7.60E-01	5.57E-01	3.32E-01
DQB1	67	32740636	2	7.60E-01	5.57E-01	3.32E-01
DQB1	28	32740753	2	7.62E-01	4.51E-01	2.30E-01
DQB1	46	32740699	2	7.80E-01	4.41E-01	2.21E-01
DQB1	47	32740696	2	7.80E-01	4.41E-01	2.21E-01
DQB1	52	32740681	2	7.80E-01	4.41E-01	2.21E-01
DPB1	205	33161588	3	8.32E-01	8.42E-01	7.78E-01
DRB1	233	32656004	3	8.44E-01	6.65E-01	6.24E-01
A	171	30019264	2	8.49E-01	7.26E-01	4.87E-01
DRB1	189	32656611	2	8.61E-01	3.88E-01	3.31E-01
DRB1	-25	32665484	3	8.80E-01	4.91E-01	4.66E-01
DRB1	73	32659929	2	8.83E-01	7.26E-01	4.03E-01
C	91	31347104	2	9.04E-01	4.07E-01	4.77E-01
DPB1	194	33161555	3	9.19E-01	9.07E-01	8.09E-01
DPB1	33	33156511	2	9.19E-01	8.77E-01	8.05E-01
DRB1	-16	32665457	3	9.40E-01	5.44E-01	5.15E-01
DQB1	-10	32742295	3	9.45E-01	4.26E-01	2.39E-01

B	307	31430883	2	9.49E-01	6.74E-01	7.40E-01
DPA1	96	33145022	2	9.57E-01	9.85E-01	7.94E-01
DPA1	190	33144526	2	9.57E-01	9.85E-01	7.94E-01
DQB1	-17	32742316	3	9.58E-01	1.20E-01	1.85E-01
DQB1	-18	32742319	3	9.61E-01	6.07E-01	5.26E-01
DQB1	9	32740810	3	9.74E-01	1.80E-01	2.73E-01
A	12	30018546	2	9.93E-01	9.50E-01	6.37E-01
B	162	31431985	2	1.00E+00	1.00E+00	1.00E+00
B	166	31431973	2	1.00E+00	1.00E+00	1.00E+00
B	211	31431264	2	1.00E+00	1.00E+00	1.00E+00
DPA1	18	33145596	2	1.00E+00	1.00E+00	1.00E+00
DPB1	91	33156685	2	1.00E+00	1.00E+00	1.00E+00

Suppl. Table 10. Stepwise analysis including the HLA-A and B the amino acid

Covariates	Risk Allele	Pvalue	OR (CI 95%)
None	HLA-B 97 Thr/Val	2.32 E-21	2.86 (2.30-3.55)
HLA-B 97 Thr/Val	HLA-A 66 Lys	1.17E-6	1.59 (1.32-1.91)
HLA-B 97 Thr/Val + HLA-A 66 Lys	HLA-B 152 Glu	0.0019	1.37 (1.13-1.68)

Suppl. Table 9. List of non-HLA variants associated with the disease

Gene	Chr	Variant	Allele	BD	Controls	P-value	OR (CI 95%)
IL23R	1	rs10889664	A	0.512	0.344	3.81E-12	2.00 (1.65-2.42)
JRKL/CNTN5	11	rs2848479	A	0.529	0.404	5.00E-08	1.65 (1.38-1.98)
IL12A	3	rs1874886	A	0.492	0.367	6.67E-08	1.67 (1.39-2.02)
PVT1	8	rs7826937	G	0.315	0.436	2.45E-07	0.60 (0.49-0.73)
PRKCG	10	rs11258609	A	0.383	0.273	5.15E-07	1.65 (1.35-2.00)
IL18R1	2	rs102372536	G	0.372	0.491	6.32E-07	0.62 (0.51-0.74)
STAT4	2	rs11694530	A	0.520	0.399	7.44E-07	1.63 (1.35-1.96)
CA11	19	rs12973839	C	0.386	0.500	1.52E-06	0.63 (0.52-0.76)
IL18R1	2	rs3732124	A	0.375	0.490	2.83E-06	0.63 (0.52-0.76)
C14orf145	14	rs2590486	A	0.378	0.487	3.52E-06	0.64 (0.53-0.78)
ATP2C2	16	rs10514605	A	0.086	0.040	4.10E-06	2.25 (1.59-3.18)
ASCC2	22	rs249398	G	0.390	0.498	4.23E-06	0.64 (0.53-0.78)
NUDFA4	7	rs4582447	C	0.540	0.429	4.54E-06	1.57 (1.29-1.89)
SLC2A13	12	rs2708494	A	0.523	0.415	5.83E-06	1.55 (1.29-1.87)
PSMA6	14	rs7152376	G	0.323	0.432	6.02E-06	0.63 (0.51-0.76)
RHPN2	19	rs11879773	A	0.537	0.432	7.34E-06	1.53 (1.27-1.85)
FUT2	19	rs602662	G	0.383	0.489	8.42E-06	0.65 (0.53-0.79)
PARK7 ERRFI1	1	rs11585541	A	0.513	0.411	9.13E-06	1.51 (1.26-1.82)
IGSF3 CD2	1	rs6677355	C	0.333	0.439	9.21E-06	0.64 (0.52-0.77)
EDG5	19	rs2288937	G	0.267	0.187	1.10E-05	1.59 (1.29-1.95)
LRP8	1	rs3737984	A	0.284	0.384	1.13E-05	0.64 (0.52-0.78)
RASIP1	19	rs2287921	A	0.381	0.485	1.17E-05	0.65 (0.54-0.79)
UBASH3A	21	rs117346108	C	0.033	0.092	1.31E-05	0.34 (0.21-0.55)
UBASH3A	21	rs7278547	G	0.349	0.453	1.54E-05	0.65 (0.53-0.79)
LOC644638 PLCL2	3	rs1375631	A	0.538	0.440	1.65E-05	1.49 (1.24-1.79)
SNAI1	20	rs6067322	G	0.290	0.389	1.75E-05	0.64 (0.53-0.79)
NCK2	2	rs4851843	G	0.452	0.355	2.04E-05	1.50 (1.24-1.81)
VKORC1	16	rs8050894	C	0.326	0.425	2.06E-05	0.65 (0.54-0.80)
CTNNA3 LOC340881	10	rs3125309	A	0.441	0.345	2.39E-05	1.50 (1.25-1.80)
ARNT2	15	rs8036233	T	0.349	0.453	2.49E-05	0.65 (0.53-0.79)
NRXN3 LOC1001316	14	rs17109461	G	0.072	0.032	2.56E-05	2.35 (1.61-3.43)
TNFRSF9 PARK7	1	rs72633102	A	0.092	0.047	2.65E-05	2.04 (1.46-2.85)
TNFRSF9	1	rs2472862	A	0.366	0.468	2.69E-05	0.66 (0.54-0.80)
LOC729159 LOC643	16	rs12598916	C	0.215	0.306	2.81E-05	0.62 (0.50-0.78)
PGM2L1	11	rs10898982	A	0.377	0.478	3.03E-05	0.66 (0.55-0.80)
FLJ42986 MAP4K4	2	rs7589943	C	0.045	0.016	3.47E-05	2.85 (1.72-4.73)
SEL1L2	20	rs6042449	A	0.510	0.407	3.60E-05	1.50 (1.24-1.81)
BCAR1 CFDP1	16	rs73605154	A	0.331	0.240	3.68E-05	1.57 (1.28-1.92)
NEDD1	12	rs4762363	A	0.158	0.237	3.79E-05	0.61 (0.48-0.77)
rs2407782	12	rs2407782	G	0.309	0.404	3.84E-05	0.66 (0.54-0.80)
NRXN3 LOC1001316	14	rs17109500	G	0.076	0.035	3.85E-05	2.26 (1.56-3.27)
NEDD1	12	rs4762361	A	0.158	0.237	3.92E-05	0.61 (0.48-0.77)
NEDD1	12	rs11108831	G	0.158	0.237	3.92E-05	0.61 (0.48-0.77)
PARK7	1	rs41278962	G	0.095	0.051	4.07E-05	1.97 (1.42-2.73)
TRAM1L1 NDST3	4	rs1350617	A	0.580	0.482	4.29E-05	1.49 (1.23-1.80)
PRNT RASSF2	20	rs968433	C	0.455	0.361	4.34E-05	1.48 (1.22-1.78)
PARK7	1	rs72634212	G	0.094	0.050	4.69E-05	1.98 (1.43-2.75)
NEDD1 LOC732096	12	rs11108828	A	0.162	0.239	4.90E-05	0.61 (0.48-0.78)
SNAI1	20	rs1467483	A	0.292	0.384	5.12E-05	0.66 (0.54-0.81)
CDK2	12	rs2069407	A	0.030	0.082	5.22E-05	0.34 (0.20-0.57)
NEDD1 LOC732096	12	rs11108824	A	0.159	0.237	5.53E-05	0.61 (0.48-0.78)
BLID UBASH3B	11	rs7944144	G	0.360	0.455	5.76E-05	0.67 (0.56-0.82)
BTBD8 C1orf146	1	rs569593	A	0.204	0.290	6.13E-05	0.63 (0.50-0.79)
CACNA1C	12	rs994901	A	0.415	0.510	6.18E-05	0.68 (0.56-0.82)

SDC4	20	rs1981431	A	0.418	0.511	6.20E-05	0.69 (0.57-0.83)
XCR1 CCR1	3	rs1491959	A	0.301	0.400	6.74E-05	0.65 (0.53-0.79)
ERAP2	5	rs17087165	G	0.363	0.458	7.57E-05	0.67 (0.55-0.82)
KIR2DL4	19	rs604999	A	0.215	0.144	7.75E-05	1.63 (1.30-2.04)
CDH18 LOC728411	5	rs10473458	C	0.399	0.491	7.96E-05	0.69 (0.57-0.83)
LOC284417 SUV420	19	rs7253685	G	0.291	0.383	8.01E-05	0.66 (0.54-0.81)
PLEK LOC554251	2	rs34852232	C	0.373	0.466	8.39E-05	0.68 (0.56-0.82)
C14orf145	14	rs8017943	A	0.314	0.403	8.41E-05	0.68 (0.55-0.83)
CRB1	1	rs10157465	C	0.044	0.100	8.41E-05	0.41 (0.27-0.64)
HERC4	10	rs597889	A	0.389	0.302	8.98E-05	1.47 (1.22-1.77)
UTS2 TNFRSF9	1	rs585672	G	0.434	0.341	9.12E-05	1.48 (1.23-1.80)
HMGCR	5	rs17238484	A	0.157	0.235	9.64E-05	0.61 (0.47-0.78)
C14orf145	14	rs327435	A	0.504	0.413	9.78E-05	1.44 (1.20-1.74)
GPR35 AQP12B	2	rs6437364	A	0.514	0.424	9.83E-05	1.43 (1.19-1.73)
GALC	14	rs10144207	A	0.470	0.381	9.99E-05	1.44 (1.19-1.74)

Resumen de resultados **y discusión**

El trabajo de esta Tesis Doctoral se ha centrado en el estudio de los factores genéticos que se encuentran involucrados en la predisposición a padecer una enfermedad con base inmunológica, la enfermedad de Behçet. El problema principal de muchos de los trabajos que abordan esta cuestión, es la inclusión de un número muy limitado de pacientes con los inconvenientes que un tamaño muestral bajo supone en términos de la limitación del poder estadístico. Este problema es debido a la baja prevalencia de la enfermedad, hecho que se acentúa en los estudios con población europea. Además, la contribución de cada uno de los genes involucrados en el desarrollo de estas patologías es relativamente baja, y este factor condiciona aún más la necesidad de cohortes más amplias para detectar asociaciones. Entre los principales éxitos de nuestro trabajo se encuentra el hecho de haber podido reclutar la cohorte más amplia de pacientes europeos existente hasta la fecha. Con esta cohorte y el diseño experimental propuesto hemos intentado proporcionar nuevos datos que contribuyan a ampliar el conocimiento del componente genético de la enfermedad. A continuación, se resumen y discuten, agrupados según los objetivos planteados, los resultados publicados en los seis trabajos que se incluyen en la presente Tesis Doctoral.

Objetivo 1. Replicación en población europea de asociaciones genes/enfermedad descritas en otras poblaciones (Trabajos 1, 2, 3 y 4).

La identificación de todos los factores genéticos que participan en la predisposición a las enfermedades complejas presenta grandes dificultades porque la contribución de cada uno de los genes involucrados en el desarrollo de estas patologías es relativamente baja. Para detectar genes que supongan un riesgo relativamente bajo, es decir, genes que presenten OR pequeñas son necesarias cohortes de población muy grandes y, como hemos comentado, los estudios realizados en BD no disponen de cohortes muy amplias (debido a su baja prevalencia), factor que determina que el poder estadístico de estos estudios sea bajo. La falta de reproducibilidad de asociaciones que encontramos frecuentemente en la literatura puede ser debida a la publicación de resultados de

asociación falsos (errores de tipo 1, falsos positivos) o a diferencias poblacionales. Por ello, para consolidar ciertas las asociaciones entre polimorfismos genéticos y enfermedad es necesario realizar estudios de replicación. Esta es la razón por la que, en el **objetivo 1** de nuestro estudio, nos propusimos replicar en población española una serie de asociaciones que habían sido descritas recientemente en otras poblaciones. Los genes que fueron seleccionados para la realización de los **trabajos 1 y 2**, son genes que se habían relacionado con la enfermedad en población asiática pero para los cuales no se habían llevado a cabo estudios de replicación en otras poblaciones. Estos genes eran además buenos candidatos desde un punto de vista funcional.

A finales de 2013, se publicaron los resultados de un GWAS realizado en población asiática (japoneses y coreanos), en el cual se propone **GIMAP** como nuevo locus de susceptibilidad para BD (Lee YJ *et al.*, 2013). En el mencionado trabajo, los autores identificaron seis SNPs asociados a la enfermedad. Estos SNPs se encuentran en un cluster génico localizado en el brazo largo del cromosoma 7 (7q36.1), que contiene genes pertenecientes a la familia **GIMAP** (**GIMAP1**, **GIMAP2**, **GIMAP4**, **GIMAP5**, **GIMAP6** y **GIMAP7**). La familia **GIMAP** es una familia de GTPasas que interviene en el desarrollo y supervivencia de los linfocitos T (Saunders A *et al.*, 2010), por lo que, variantes de los genes que codifican estas moléculas, podrían estar relacionadas con la predisposición a padecer enfermedades autoinflamatorias. Sin embargo, la asociación de **GIMAP** con BD todavía no había sido replicada en otras poblaciones y, por tanto, no estaba confirmada.

Según los resultados de nuestro estudio, no se observaron diferencias estadísticamente significativas en la distribución de las frecuencias entre pacientes y controles para ninguno de los seis SNPs que se encontraron asociados en el trabajo de Lee YJ *et al.* (**trabajo 1, tabla 1**). No podemos descartar la existencia de variaciones poblacionales que expliquen la disparidad de resultados, sin embargo, estas discrepancias también pueden ser debidas a errores tipo I propiciados, por ejemplo, por la existencia de algún tipo de sesgo en la selección

de la población control. En este sentido, para tratar de explicar las razones de las discrepancias entre nuestro trabajo y el anteriormente publicado, se recopilamos las frecuencias descritas para estos SNPs en el proyecto 1000 Genomas considerando las poblaciones IBS, JPT y ASN como poblaciones de referencia para la española, japonesa y coreana, respectivamente. Como se puede observar en la **tabla 4**, que resume esta información, las frecuencias encontradas en nuestro trabajo en población control para el alelo de riesgo son similares a las descritas para población IBS; mientras que, las frecuencias descritas para las poblaciones control coreanas y japonesas para el alelo de riesgo, en el trabajo de Lee YJ *et al.*, son inferiores a las publicadas para JPT y ASN en el proyecto 1000 Genomas. Dicho de otra forma, las frecuencias de las poblaciones de referencia de poblaciones asiáticas se asemejan más a las de las cohortes de pacientes incluidas en el trabajo de Lee YJ *et al.* que a las de sus controles. Hay que tener en cuenta que la existencia de variación de las frecuencias de los alelos dentro de las poblaciones es una fuente de identificación de asociaciones con errores tipo I (Cardon LR *et al.*, 2003; Ziv E *et al.*, 2003). En este sentido, aunque se ha visto que existen gradientes poblacionales dentro de Europa (Heath SC *et al.*, 2008; Lao O *et al.*, 2008), estos gradientes son mucho menos acusados que en otras poblaciones, como la población asiática (Rosenberg NA *et al.*, 2002; The HUGO Pan-Asian SNP Consortium, 2009; Reich D *et al.*, 2009). Este es el motivo por el que es muy importante comprobar que la población de estudio está bien seleccionada y que no existe estratificación para evitar sesgos y falsos positivos (Tiwari HK *et al.*, 2008).

Tabla 4. Frecuencias de los 6 SNPs de los genes *GIMAP* estudiados en diferentes poblaciones

Gen	SNP	Alelo riesgo	Población	Frecuencia alelo riesgo		
				Casos	Controles	Población de referencia
GIMAP 4	rs1916012	A	Coreana	0.58	0.48	0.52
			Japonesa	0.47	0.41	0.49
			Española	0.57	0.58	0.58
	rs1522596	A	Coreana	0.59	0.48	0.52
			Japonesa	0.47	0.41	0.49
			Española	0.58	0.58	0.59
	rs1608157	C	Coreana	0.59	0.48	0.52
			Japonesa	0.47	0.41	0.49
			Española	0.56	0.58	0.58
	rs1026606 ₉	A	Coreana	0.49	0.41	0.51
			Japonesa	0.46	0.40	0.46
			Española	0.51	0.49	0.51
GIMAP 2	rs1025648 ₂	T	Coreana	0.50	0.43	0.49
			Japonesa	0.46	0.40	0.46
			Española	0.60	0.58	0.60
	rs2286900	A	Coreana	0.21	0.14	0.21
			Japonesa	0.18	0.20	0.21
			Española	0.10	0.09	0.07

Las poblaciones de referencia del proyecto 1000 Genome Project son ASN, JPT y IBS, para coreanos, japoneses y españoles respectivamente.

TNFAIP3 y *JAK1* han sido recientemente identificados en dos trabajos independientes como nuevos locus de susceptibilidad con BD en una población china, concretamente en población Han (Hou S *et al.*, 2013; Li H *et al.*, 2013). *TNFAIP3* (6q23) codifica una ubiquitinasa involucrada en la respuesta a través de NF-kappa-B (Boone DL *et al.*, 2004) mientras que, *JAK1* (1p32.3-31.3) interviene en las rutas de señalización de los IFN α/β y γ (Babon JJ *et al.*, 2014). Diversas

razones nos impulsaron a diseñar un estudio de replicación para confirmar la asociación de estos genes en nuestra población. La primera, como en el caso anterior, era que la asociación no se había identificado en otras poblaciones, ya que, no se había hecho ningún estudio de replicación y tampoco se había descrito en ninguno de los GWAS realizados hasta la fecha (Fei Y *et al.*, 2009; Remmers EF *et al.*, 2010; Mizuki N *et al.*, 2010; Hou S *et al.*, 2012; Lee YJ *et al.*, 2013). Otra razón era que el 100% de los individuos de la población de pacientes incluida en los 2 trabajos previos padecían uveítis y, por tanto, no se sabía si, en caso de confirmarse, la asociación estaba relacionada con la predisposición a la enfermedad o con el desarrollo de una característica clínica determinada, concretamente con la uveítis. Además, desde un punto de vista funcional, ambos genes son buenos candidatos y se ha descrito asociación con diferentes enfermedades de base inmunológica (Hu K *et al.*, 2013). *TNFAIP3* se ha descrito como relacionado con SSc en población caucásica (Dieudé P *et al.*, 2010), concretamente un SNP, rs5029939, que está en LD ($r^2=0.88$ y $D'=1$) con el rs9494885, incluido en nuestro estudio. También se ha descrito asociación de otros 2 SNPs: rs10499194 y rs2230926 con RA en población china (Zhang X *et al.*, 2014). En nuestro trabajo incluimos el primero pero no incluimos el segundo porque su MAF en población caucásica es muy baja (0.02). En cuanto a *JAK1* hay un estudio en población china que encuentra asociación con 3 SNPs, con elevado LD entre sí, con el síndrome Vogt-Koyanagi-Harada (VKH). En nuestro trabajo incluimos 1 de estos 3 SNPs.

Los resultados obtenidos en nuestra población no muestran ninguna diferencia en la distribución de las frecuencias entre casos y controles para ninguno de los 5 SNPs analizados (**trabajo 2, tabla 1**). Como en el caso anterior, analizamos si la disparidad de resultados podría atribuirse a un sesgo originado, por ejemplo, por la selección de la población control. Sin embargo, al contrario que en el caso anterior y como se puede apreciar en la **tabla 5** las frecuencias de las variantes de riesgo en las poblaciones control son similares a las descritas en

las poblaciones de referencia del proyecto 1000 Genomas, siendo CHB la población de referencia elegida para la población Han.

Tabla 5. Frecuencias de los cinco SNPs estudiados en los genes *TNFAIP3* y *JAK1* en diferentes poblaciones.

Gen	SNP	Alelo de riesgo	Población	Frecuencias alelo de riesgo		
				Casos	Controles	Población de referencia
<i>TNFAIP3</i>	rs10499194	T	Chinos Han	0.02	0.04	0.02
			Española	0.32	0.30	0.35
	rs610604	G	Chinos Han	0.09	0.07	0.12
			Española	0.36	0.33	0.35
	rs9494885	C	Chinos Han	0.17	0.10	0.08
			Española	0.11	0.12	0.08
<i>JAK1</i>	rs2780815	G	Chinos Han	0.89	0.85	0.85
			Española	0.49	0.47	0.50
	rs310241	G	Chinos Han	0.33	0.27	0.27
			Española	0.27	0.28	0.29

Las poblaciones de referencia del proyecto 1000 Genome Project son CHB y IBS para chinos Han y españoles respectivamente.

Como se ha comentado anteriormente, nos planteamos la posibilidad de que la asociación de estos genes estuviera relacionada con la presencia de uveítis, ya que, en la cohorte de descubrimiento, todos los pacientes presentaban esta manifestación clínica. También exploramos la posibilidad de que la asociación estuviera relacionada con la presencia de otros factores de riesgo genético, por lo que se investigó la asociación de estos factores en subgrupos de pacientes estratificados según la presencia de HLA-B*51. En ningún caso se detectaron diferencias significativas en la distribución de estos marcadores (**artículo 2, tabla**

2). Por tanto, los estudios realizados hasta el momento en los genes de *TNFAIP3* y *JAK1* sugieren asociación con BD en población asiática, pero no en población europea. Este sería un ejemplo más de diferencia entre poblaciones como se ha descrito en el caso del gen *PADI4* y su asociación con RA (Martinez A *et al.*, 2005; Suzuki A *et al.*, 2003)

En resumen, los resultados obtenidos en población Española no reproducen las asociaciones descritas en población asiática. Esta enfermedad, al igual que la mayoría de las enfermedades complejas, se caracteriza por ser una enfermedad muy heterogénea, por lo que las discrepancias que se observan podrían ser debidas a una asociación específica con población asiática y no con población caucásica. Otra posible explicación a la falta de consistencia en las replicaciones podría ser la existencia de una distribución diferente de las manifestaciones clínicas en las diferentes poblaciones de estudio, debido al carácter sistémico y multifactorial de la enfermedad. En esta línea debemos hacer notar que, como se ha comentado en el caso de los trabajos de *JAK1* y *TNFAIP3* el 100% de los pacientes muestra manifestaciones oculares (Hou S *et al.*, 2013; Li H *et al.*, 2013). Por tanto aunque en nuestro estudio, no se encontraron evidencias de asociación con ésta y otras manifestaciones clínicas, hay que tener en cuenta la limitación en el poder estadístico que origina la disminución del tamaño muestral por la estratificación en diferentes subgrupos. Sin embargo, en el trabajo donde se describe la asociación con los genes *GIMAP*, la cohorte de descubrimiento incluía un porcentaje de individuos con uveítis similar al de nuestra cohorte (43.5% y 49.8% respectivamente), aunque en la cohorte de replicación (también asiática) incluida en ese trabajo tenía un porcentaje mucho más elevado de pacientes con esta manifestación clínica (84.6%) (Lee YJ *et al.*, 2013). Los resultados publicados por nuestro grupo en los trabajos 1 y 2 sugieren que ninguno de estos genes (*GIMAP*, *JAK1* y *TNFAIP3*) se encuentran asociados a BD en población española, aunque no se puede descartar totalmente una posible influencia en las características clínicas.

CCR5 y *PTPN22* son genes reiteradamente estudiados porque las moléculas que codifican desempeñan un papel muy importante en diversas enfermedades inmuno-mediadas, además, en ambos casos se conoce la variante funcional responsable de las asociaciones.

El polimorfismo más ampliamente estudiado del gen *CCR5*, no es un SNP, sino una delección de 32 pares de bases en el gen *CCR5* (*CCR5*Δ32). Esta delección tiene como resultado un cambio de marco de lectura que origina la aparición de un codón de parada prematuro, por lo que, la variante se traduce en una proteína truncada incapaz de ejercer su función correctamente (Charo IF *et al.*, 2006). Cuatro publicaciones previas, que incorporan datos de seis poblaciones diferentes (UK, turcos, palestinos, iraníes, italianos y portugueses) han estudiado la asociación de esta variante con BD (Yang X *et al.*, 2003; Mojtabehi Z *et al.*, 2006; Atzeni F *et al.*, 2011; Bettencourt A *et al.*, 2014). Los resultados de nuestro estudio, que concuerdan con 3 de los estudios previos, sugieren que la variante *CCR5*Δ32 no se asocia con la predisposición a la enfermedad, ni tampoco con las características clínicas ni con la presencia de HLA-B*51. De hecho, solo el estudio realizado en población italiana sugiere que *CCR5*Δ32 sea un factor de riesgo en la predisposición a padecer la enfermedad (Atzeni F *et al.*, 2011). Debido a que el poder estadístico de todos los estudios individuales es bajo tanto por el escaso tamaño muestral como por la baja frecuencia de la variante en todo el mundo, nos planteamos realizar un meta-análisis agrupando todos los datos disponibles hasta el momento. La **tabla 2 del trabajo 3**, muestra un resumen de las características de todos los estudios que se han incorporado en el meta-análisis así como el poder estadístico de cada uno de ellos y el global. El resultado de este meta-análisis que incluía un total de 1114 pacientes y 1786 controles sanos, no evidencia asociación de la variante de *CCR5* con la enfermedad, por lo que podemos descartar que *CCR5*Δ32 tenga un papel importante en esta patología, ya que el poder estadístico global es adecuado, 93%, para detectar asociación con $OR \geq 1.5$

En cuanto al gen *PTPN22*, se ha descrito que la variante funcional, R620W, tiene un papel importante en múltiples patologías inmuno-mediadas, como RA y SSc entre otras (Begovich AB *et al.*, 2004; Dieudé P *et al.*, 2008). Con respecto a BD existen 2 trabajos previos que investigan la asociación de esta variante con la enfermedad en 3 poblaciones diferentes (ingleses, turcos y medio oriente), sin embargo, los resultados de estos trabajos no son concluyentes (Baranathan V *et al.*, 2007; Sahin N *et al.*, 2007). De hecho, de las 3 poblaciones estudiadas, se encontró asociación en 2, ingleses y turcos, mientras que, para la población de oriente medio, las diferencias no fueron significativas. Nuestros resultados tampoco evidencian asociación de esta variante con la enfermedad. De hecho, la distribución de la frecuencia de la variante en pacientes y controles de población española es prácticamente igual y muy similar a la que se describe en el proyecto 1000 Genomas para poblaciones IBS y TSI. Otra cuestión a resaltar, es que en las poblaciones en las que se ha descrito asociación de R620W con BD (ingleses y turcos) lo que se describe es que la variante minoritaria confiere protección a la enfermedad (OR=0.46, OR=0.65, respectivamente) (Baranathan V *et al.*, 2007; Sahin N *et al.*, 2007). Sin embargo, en la inmensa mayoría de enfermedades en las que se ha estudiado este polimorfismo, o no se encuentra asociación (psoriasis, SSc, UC, etc.) o la variante minoritaria confiere riesgo (GCA OR=1.62, RA OR=1.65, SLE OR=1.46, enfermedad de Graves OR=1.5, vitíligo OR=1.98, etc) (Stanford SM *et al.*, 2014). A pesar de que la relación entre esta variante y la enfermedad no está bien establecida, en revisiones de *PTPN22* se da por cierta la asociación de R620W con BD (Stanford SM *et al.*, 2014). Con el fin de aportar nuevos datos y teniendo en cuenta que los condicionamientos de los estudios individuales son similares a los que se exponían en el caso de *CCR5* nos planteamos realizar un meta-análisis agrupando la información disponible con respecto a esta variante en BD. La **tabla 2 del trabajo 4** resume la información de las características de los estudios incluidos en este meta-análisis donde se agrupaban datos de cuatro poblaciones con un total de 794 pacientes y 1876 controles sanos. Como resultado de este análisis, no se observó asociación entre la variante R620W de *PTPN22* y BD.

Sin embargo, como se ha descrito que *PTPN22* desempeña un papel clave en muchas enfermedades inmuno-mediadas, nos planteamos la posibilidad de que la asociación del gen con BD residiera en otra variante diferente de R620W, ya que, esta variante es prácticamente inexistente en países con una alta prevalencia de la enfermedad como son los países asiáticos y Turquía. Por este motivo decidimos estudiar la variabilidad genética en todo el gen, genotipando un total de 6 SNPs (5 tSNPs y otro SNP que se ha descrito que es una variante funcional) con los que se cubría el 100% de la región génica. Nuestros resultados no demuestran diferencias estadísticamente significativas para ninguno de estos polimorfismos entre los grupos de pacientes y controles. En concordancia con nuestro trabajo se encuentran los resultados de un estudio en población china Han donde se investigó la asociación de 3 polimorfismos en *PTPN22* entre los que no estaba incluido R620W (Zhang Q *et al.*, 2012). En el mencionado trabajo, con un tamaño poblacional adecuado, no se encuentran diferencias en la distribución de las frecuencias entre los grupos de pacientes y controles para ninguno de los polimorfismos incluidos. Además, no se ha descrito asociación del gen con la enfermedad en ninguno de los GWAS realizados hasta la fecha (Fei Y *et al.*, 2009; Remmers EF *et al.*, 2010; Mizuki N *et al.*, 2010; Hou S *et al.*, 2012; Lee YJ *et al.*, 2013). Es más, en uno de los GWAS realizado en población turca, los autores realizaron una búsqueda dirigida de los genes con mayor relevancia en este tipo de enfermedades entre los que incluyeron *PTPN22*. Ninguno de los SNPs de este gen incluidos en el GWAS alcanzaba significación estadística y el valor más bajo de p para cualquiera de ellos era de 0.27 (Remmers EF *et al.*, 2010). Por tanto, nuestros resultados junto con la mayoría de los publicados sugieren que este gen no desempeña un papel importante en la predisposición a padecer esta enfermedad. Como se ha explicado, la variante R620W desempeña un papel muy importante en otras patologías de base inmunológica y, por ello, hay muchos trabajos que investigan esta variante. Se ha comprobado que en los casos de verdadera asociación, los resultados son altamente reproducibles, encontrándose relación con una determinada patología en muchas poblaciones. Además, parece que la

asociación de esta variante está bastante restringida a enfermedades en los que la presencia de autoanticuerpos tiene un papel relevante (Burn GL *et al.*, 2011).

En resumen, los resultados de nuestros trabajos nos indican que podemos descartar que tanto la delección $\Delta 32$ de *CCR5* como el gen *PTPN22* desempeñen un papel importante en la predisposición a la enfermedad. En ambos casos, nuestro trabajo incluye la cohorte con el número más elevado de casos y controles y con el mayor poder estadístico de todos los estudios publicados hasta la fecha. Además, la metodología utilizada por nuestro grupo es más reproducible que la usada en otros estudios. Y por último, en algunas áreas geográficas donde la enfermedad es relativamente prevalente, las variantes de riesgo para *CCR5* y *PTPN22* son prácticamente inexistentes, lo que sugiere que estas variantes no deben de desempeñar un papel importante en la patología. Sabemos que existe un fuerte sesgo de publicación de estudios de asociación genética que presentan hallazgos positivos (Ionnidis JP *et al.*, 2001) por lo que los resultados de este tipo de estudios deben de ser tratados con prudencia hasta que se validen en estudios de replicación.

Los trabajos discutidos en este objetivo son ejemplos de la importancia de los estudios de replicación ya que los resultados en un trabajo individual o en una sola población no son suficientes para dar por establecida la asociación con una determinada patología.

Objetivo 2. Búsqueda de nuevos genes candidatos (trabajos 5 y 6).

La región HLA de clase I es la región del genoma más sólidamente asociada con BD, sin embargo, solo explica el 20-30% de la carga genética de la enfermedad (Yazici H *et al.*, 2007). Por tanto, fuera de esta región debe haber otros genes que condicionen la predisposición a esta enfermedad y de hecho, recientemente, se han descrito otros genes, como *IL23R* e *IL10*, que han sido consistentemente asociados con la enfermedad en diferentes poblaciones, ya que,

los estudios de replicación proporcionan resultados homogéneos (Mizuki N *et al.*, 2010; Remmers EF *et al.*, 2010; Montes-Cano MA *et al.*, 2013; Carapito R *et al.*, 2015; Xavier JM *et al.*, 2012; Wu Z *et al.*, 2014; Kang EH *et al.*, 2014; Matos M *et al.*, 2014; Jiang Z *et al.*, 2010; Kim ES *et al.*, 2012; Yalçın B *et al.*, 2014). En los últimos años se han realizado varios estudios GWAS (“Introducción”, tabla 3), que han aportado nueva información, pero aun así, los logros conseguidos por estos estudios quedan lejos de los logros conseguidos en otras patologías en las que el aporte de los GWAS ha sido clave en el incremento del conocimiento de las bases de la enfermedad. Con respecto a BD, queda por conocer gran parte del componente genético, por lo que en este objetivo nos planteamos llevar a cabo una búsqueda de nuevos genes que se encuentren involucrados en dicha patología. Para alcanzar este objetivo nos planteamos dos estrategias diferentes para cada una de las cuales elegimos diferentes plataformas de genotipado como se explica en el apartado “Hipótesis, Objetivos y Diseño Experimental” dependiendo, en este caso, del número de SNPs a incluir en cada estudio.

Como se ha explicado en la introducción, la BD se clasifica como una enfermedad de patrón mixto, ya que, se considera que tanto la inmunidad adaptativa como la inmunidad innata se encuentran involucradas en la patogenia. Sin embargo, faltan evidencias que sustenten el papel de la inmunidad innata, lo que nos hace pensar que parte de la carga genética que no se conoce todavía esté relacionada con este brazo del sistema inmunitario. Teniendo en cuenta que uno de los posibles mecanismos etiopatológicos de la enfermedad, está relacionado con que determinados agentes infecciosos sean los desencadenantes de la misma en individuos genéticamente predispuestos, nos planteamos estudiar si determinados receptores de la inmunidad innata podrían estar relacionados con la enfermedad. En este sentido, los receptores de reconocimiento de patrones son los sensores de la inmunidad innata encargados de reconocer a los patógenos y por tanto, estos sensores y concretamente los de ácidos nucleicos, podrían desempeñar un papel en BD. Este fue el objetivo que nos planteamos en el **trabajo 5** para el que se seleccionaron un total de 9 genes candidatos. El esquema de la figura 9 representa

los 9 receptores elegidos, sus ligandos y las rutas que se activan tras el reconocimiento. El diseño final incorporaba 74 tSNPs que nos permitirían capturar la información correspondiente a 216 SNPs (**trabajo 5, tabla S1**) por lo que elegimos una plataforma de genotipado de medio rendimiento.

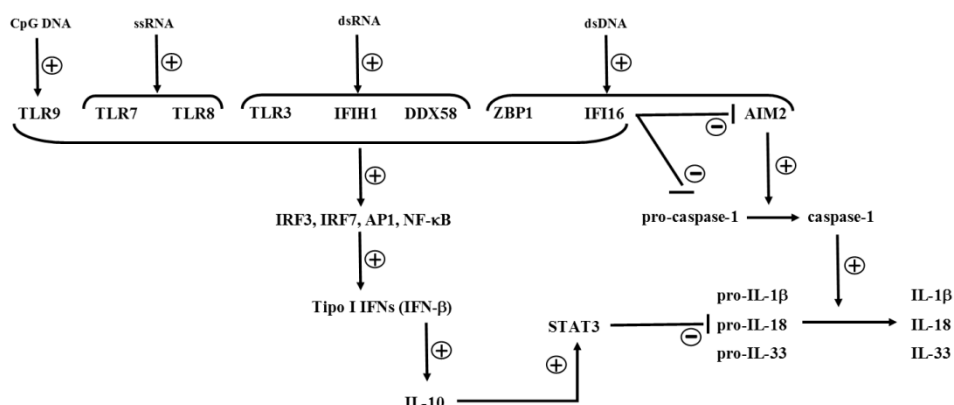


Figura 9. Receptores de reconocimiento de patrones involucrados en el reconocimiento de ácidos nucleicos incluidos en el estudio. En el esquema se representan los 9 genes incluidos en el estudio, las dianas que reconocen los receptores y las rutas que se desencadenan tras el reconocimiento de dichas dianas.

Como resultado de este estudio identificamos el cluster genético ***IFI16/AIM2*** como un nuevo locus de susceptibilidad a BD en población española. *IFI16* y *AIM2* se encuentran juntos formando un cluster de 70 kb en el cromosoma 1 (1q25.2) y se transcriben en direcciones opuestas. Como se puede ver en la **tabla 3 del artículo 5**, que muestra el análisis de haplotipos constituidos por los SNPs que se encontraron asociados, encontramos 3 haplotipos diferentes en nuestra población, uno de ellos confiere riesgo ($p=0.015$; OR (IC95%)=1.41 (1.06-1.86)), otro protección ($p=0.009$; OR (IC95%)=0.65 (0.47-0.90)) y otro lo consideramos neutro porque no se encontraron diferencias en la distribución de frecuencias entre pacientes y controles. Estos dos genes pertenecen a una misma familia de proteínas denominada Ifi-200 cuyo nombre hace referencia a que son inducibles por IFN, además, ambas proteínas son citoplasmáticas y reconocen dsDNA. La

figura 10 representa un esquema de las rutas en las que se encuentran involucradas ambas proteínas.

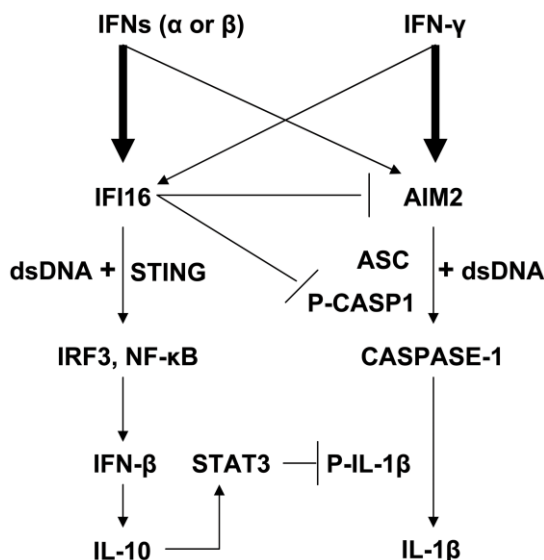


Figura 10. Cascada de señales en las que intervienen las proteínas IFI16 e AIM2. Aunque ambas reconocen dsDNA, los efectos que desencadenan tras su activación son opuestos. AIM2 forma parte de un inflamasoma, al reclutar a la proteína adaptadora ASC y a la caspasa 1, este complejo va a dar lugar a las formas activas de IL-1 β e IL-18. Así como el efecto de AIM2 es pro-inflamatorio, el de IFI16 va a ser anti-inflamatorio. Por un lado el reconocimiento de dsDNA va a dar lugar a la síntesis de IL-10 (interleucina anti-inflamatoria). Además, IFI-16 puede inhibir la ruta pro-inflamatoria inhibiendo directamente al inflamasoma o bien al formar heterodímeros con AIM2 mediante los dominios PYD. (Veeranki S *et al.* PLoS ONE, 2011)

La proteína AIM2 forma parte del inflamasoma y su activación produce un aumento de proteínas proinflamatorias. En cambio, IFI-16 tiene una función contraria, ya que, es un regulador negativo directo del inflamasoma, y además activa una ruta que finaliza con el aumento de interleucinas anti-inflamatorias (Veeranki S *et al.*, 2011). Se ha demostrado que los niveles de expresión de ambas

proteínas son inversos (Duan X *et al.*, 2011), por tanto, el ratio de proteínas AIM2/IFI16 existente en la célula determina el tipo de respuesta (pro o anti-inflamatoria) que se va a desencadenar tras el reconocimiento de dsDNA.

Hay estudios que sugieren que BD es una enfermedad relacionada con el inflammasoma. En este sentido, se ha descrito un aumento de los niveles de interleucinas inflamatorias (IL1 β , IL18 e IL33) en suero de los pacientes con respecto a la población sana (Düzgün N *et al.*, 2005; Musabak U *et al.*, 2006; Hamzaoui K *et al.*, 2013). Además, se ha visto que los pacientes con BD responden bien a tratamientos que bloquean IL1 β (Gül A *et al.*, 2012). También en este sentido, a veces es difícil establecer un diagnóstico diferencial con enfermedades atoinflamatorias típicas que cursan con fiebres recurrentes como la fiebre mediterránea familiar, que es una enfermedad monogénica causada por mutaciones en el gen *MEFV*. Otra patología autoinflamatoria si bien poligénica es la CD que se ha relacionado con mutaciones en el gen *NOD2*. En esta línea, hay estudios que sugieren asociación de los genes *MEFV* y *NOD2* con BD. La implicación de estos genes, sustentaría que la respuesta inmune innata contra los componentes bacterianos estaría relacionada la patogénesis de la BD (Kirino Y *et al.*, 2013; Tasliyurt T *et al.*, 2013). Por estos motivos decidimos llevar a cabo un estudio en el que determinamos los niveles de expresión de mRNA de *IFI16* y *AIM2* en individuos que portaban el haplotipo de riesgo y los que portaban el haplotipo de protección usando como referencia aquellos individuos que eran homocigotos para el haplotipo mayoritario, que corresponde con el haplotipo neutro. Este ensayo se llevó a cabo en individuos sanos para evitar sesgos derivados de la variedad de tratamientos a las que pueden ser sometidos los pacientes. No se encontraron diferencias significativas en los niveles de expresión de *AIM2* entre ninguno de los grupos de individuos. Sin embargo, se encontraron diferencias estadísticamente significativas en los niveles de *IFI16* entre los portadores del haplotipo de riesgo y los portadores del haplotipo de protección ($p=0.022$). La media de expresión de los individuos que portaban el haplotipo de protección era la mayor (1.23 ± 0.50) y la de los individuos con el haplotipo de

riesgo era la menor (0.99 ± 0.29), mientras que los de referencia muestran niveles intermedios (1.08 ± 0.46) (**trabajo 5, tabla 4**). Estos resultados determinan también diferencias significativas en el ratio de AIM2/IFI16 entre estos grupos. Los individuos con el haplotipo de riesgo presentan un ratio AIM2/IFI16 más elevado (1.11 ± 0.58), existen diferencias significativas cuando comparamos este ratio con los individuos que portan el haplotipo de protección (0.84 ± 0.46 , $p=0.049$) y con los individuos con el haplotipo neutro (0.81 ± 0.35 , $p=0.022$). El ratio de proteínas AIM2/IFI16 existentes en la célula podría ser el responsable del balance entre las respuestas pro-inflamatorias y anti-inflamatoria tras el reconocimiento de dsDNA y, nuestros resultados sugieren que los individuos que portan el haplotipo que hemos considerado de riesgo tendrían mayor tendencia hacia una respuesta pro-inflamatoria por un menor cociente en los niveles de expresión AIM2/IFI16.

En resumen, nuestros resultados sugieren relación entre el cluster *AIM2/IFI16* y la enfermedad de BD y también, que esta asociación podría ser explicada por los niveles de expresión basales de *IFI16*. Aun así, debido a que es la primera vez que se identifica asociación entre estos genes y la enfermedad, la consistencia de esta relación no se demostrará hasta que no se realicen estudios de replicación en otras cohortes y poblaciones.

Como segunda estrategia para alcanzar el **objetivo 2**, nos planteamos utilizar una de las plataformas de genotipado con la que más éxitos se han obtenido en los últimos años a la hora de identificar nuevas variantes de riesgo en enfermedades de base inmunológica, el iChip. Aunque en el momento en que comenzamos este estudio ya se habían publicado varios GWAS con pacientes de diferentes poblaciones en esta enfermedad, estos estudios se realizaron con cohortes turcas, de medio oriente y asiáticas en su mayoría, pero con muy poca participación de europeos (de hecho, sólo se incluyen europeos en la cohorte de replicación de un GWAS) y con la inclusión de un número relativamente bajo de individuos en la mayoría de los casos (**introducción, tabla 3**)

A pesar de que tanto los GWAS como el iChip son estudios de alto rendimiento, tal y como se explicó en el capítulo “Introducción”, los chips usados para ambos estudios están diseñados siguiendo diferentes estrategias, por lo que cada uno cubre diferentes objetivos. Hasta la fecha, aparte de nuestro trabajo, solo se han publicado datos parciales de resultados de iChip en BD, concretamente sólo los datos correspondientes a la región HLA (Hughes T *et al.*, 2013). Por tanto, este trabajo es el primero en aportar datos completos obtenidos a través de la plataforma iChip en BD.

En general en todo trabajo, y en particular los estudios de alto rendimiento, el análisis de los datos obtenidos del genotipado deben someterse a estrictos controles de calidad antes de validarlos y pasar a la fase de análisis y búsqueda de asociaciones. Los criterios de calidad usados en nuestro estudio se describen en profundidad en el **trabajo 6**, pero cabe destacar que, una de las características del iChip es que incorpora muy buenos marcadores de ancestría para población europea, de manera que, se puede comprobar de forma muy fiable si la población de estudio es homogénea y si los controles se han seleccionado adecuadamente. Además, la región HLA está densamente cubierta, y esto nos permite llevar a cabo un análisis de parentesco, que evite información redundante por la presencia de familiares en el estudio. Después de aplicar los criterios de calidad se eliminaron 9 pacientes y 63.506 SNP y fueron aptos para analizar 277 casos, 1517 controles (usamos una cohorte de controles, previamente publicada en Carmona FD *et al.*, 2015) y 133.018 SNPs.

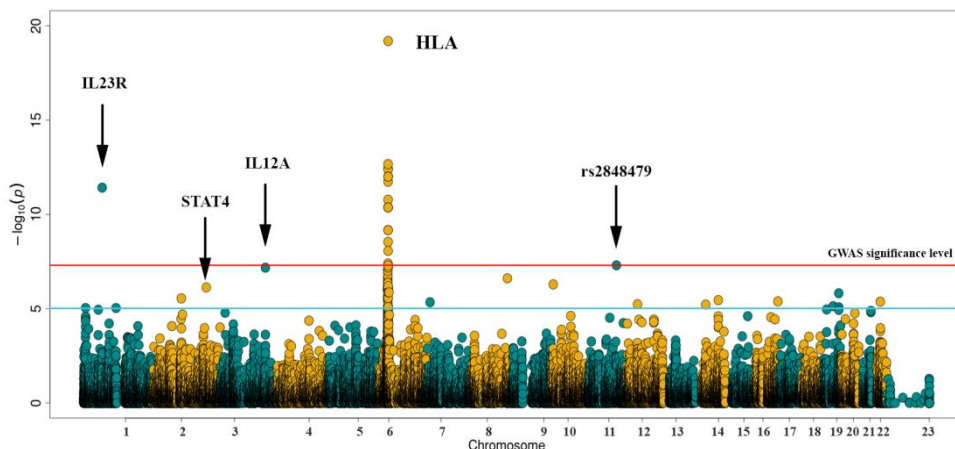


Figura 11. Manhattan plot de los resultados del iChip. El Manhattan plot es una representación gráfica del estudio de asociación de todas las variantes analizadas en este chip. Cada círculo corresponde con una variante incluida en el estudio, el eje de la X indica la posición de las variantes en el genoma y el eje Y representa el logaritmo del valor de p para cada variante. El umbral de asociación establecido para los estudios GWAS es $p=5 \times 10^{-8}$ se representa en la figura por una línea roja. Aquellas variantes que presentan valores de p inferiores al umbral se consideran asociadas. La región que se encuentra limitada entre la línea roja y la azul incluye variantes con valores de p de 5×10^{-4} a 5×10^{-8} , este área se considera “zona gris” y las señales que se localizan en ella se consideran posiblemente asociadas.

Como era de esperar y se puede comprobar en el Manhattan plot (**trabajo 6, figura 1**), la región con asociación más fuerte en nuestra población fue la región HLA; los resultados de este trabajo en esta región se discuten en el objetivo 3. Fuera de la región HLA, la única señal que supera el umbral de significación corresponde con *IL23R*, dato concordante con los publicados previamente. De hecho, nuestro grupo ya había descrito asociación de *IL23R* en nuestra población en un trabajo de replicación previamente publicado (Montes-Cano MA *et al.*, 2013). Existen controversias en cuanto que al uso estricto del valor $p=5 \times 10^{-8}$, considerado como umbral en los GWAS, en los estudios con iChip. Como se ha dicho, la estrategia y el diseño de ambas plataformas es diferente, de hecho, en iChip se estudian menos señales independientes (lo que significaría menor necesidad de correcciones estadísticas) que en los estudios GWAS por lo que hay opiniones que defienden que el umbral de significación debería ser menos restrictivo (Sham PC *et al.*, 2014). Sin embargo, la mayoría de los estudios

publicados hasta la fecha con esta plataforma y diferentes patologías mantienen el umbral que se estableció para los estudios GWAS. El tamaño muestral de nuestro estudio limita su poder estadístico porque es difícil de llegar a este nivel de significación con este tamaño. Alcanzar valores de *p* tan elevados es muy difícil con las cohortes de BD existentes hasta la fecha. De hecho, *IL23R* solo supera ese valor en uno de los GWAS realizados (Hou S *et al.*, 2012), en otro, la señal más alta se encuentra en la zona gris (Remmers EF *et al.*, 2010). Como hemos comentado, *IL23R* se ha replicado en muchas poblaciones y, de hecho, después de HLA-B*51 se le considera el gen más fuertemente asociado con la enfermedad. Este caso es un ejemplo útil para discutir el valor umbral de *p* en los estudios a gran escala. Que una señal supere el umbral nos indica que es muy improbable que las diferencias entre las frecuencias que se hayan encontrado se deban al azar. Sin embargo, el resto de señales que no superan el umbral pueden ser verdaderos positivos. Por ello que se establece una “zona gris”, en la que encontraremos señales que con bastante probabilidad estén asociadas, pero que sean necesarios estudios de replicación para validarlas.

En la **tabla S11 del trabajo 6** se muestran las señales que se encontraron en la zona gris, de todas ellas seleccionamos para replicación en otra cohorte de pacientes y controles de nuestra población las siguientes: el rs2848479 (*JRKL/CNTN5*), rs1874886 y rs17810546 (*IL12A*), rs11694530 (*STAT4*), rs604999 (*KIR2DL4*), rs7944144 (*UBASH3*) y rs7616215 (*CCR1*). Estas señales fueron elegidas por diferentes razones, el valor de *p* para la asociación del rs2848479 se encuentra justo en el umbral, es decir, es la señal más fuerte de todas las de la zona gris, si bien, no hay ninguna razón funcional conocida para su inclusión, ya que, se encuentra situado entre 2 genes que no han sido relacionados con ninguna enfermedad inmuno-mediada. Las señales localizadas en *IL12A*, *STAT4*, *UBASH3B*, *KIR2DL4* y *CCR1* se seleccionaron porque estos genes se habían relacionado con la enfermedad en poblaciones no europeas (Kirino Y *et al.*, 2013; Kappan JH *et al.*, 2015; Hou S *et al.*, 2012, Fei Y *et al.*, 2009, Kim SJ *et al.*, 2013). De hecho, genotipamos también rs17810546 (*IL12A*) en la cohorte

completa porque era el SNP que se había encontrado asociado en estudios anteriores. Los resultados del estudio de replicación demuestran diferencias significativas en la segunda cohorte de pacientes y controles para todos los SNPs salvo para los correspondientes a los genes *KIR2DL4* y *UBASH3B* incluso en estos casos, la tendencia en la segunda cohorte fue la misma que en la primera por lo que, en todos los casos, las diferencias fueron significativas en la cohorte conjunta.

En resumen, los resultados de nuestro estudio sugieren un nuevo loci de susceptibilidad en esta enfermedad, *JRKL/CNTN*. Estos resultados deben ser reproducidos en otras poblaciones para confirmar la asociación. Además, encontramos asociación con SNPs de los genes *IL23R*, *IL12A*, *STAT4*, *UBASH3B*, *KIR2DL4* y *CCR1* que habían sido previamente descritas en otras poblaciones, por lo que, nuestro trabajo se puede considerar una replicación de los trabajos previos y sugiere que dichos genes pueden desempeñar un papel importante en la patología de BD. Todos estos genes juegan un papel importante en la modulación de la respuesta inmune. Así, *IL23R* codifica una subunidad del receptor de la IL23 que se expresa en la superficie de las células Th17 y de los macrófagos. El complejo IL23/IL23R promueve la polarización de la respuesta T hacia Th17 (Iwakura Y *et al.*, 2006). En cuanto a *IL12A*, codifica una subunidad del heterodímero que forma IL12. Esta citocina juega un papel importante en la polarización de la respuesta inmune hacia Th1 y en la producción de IFN γ por los linfocitos T y por las células NK por lo que está relacionada con la producción de citocinas proinflamatorias (Chang JT *et al.*, 1999). Con respecto a *STAT4*, es un factor de transcripción activado por citocinas como la IL12 y IL23 que está involucrado en la diferenciación de los linfocitos hacia Th1 y Th17 (Morinobu A *et al.*, 2002). *CCR1* codifica un receptor de quimioquinas que, mediante la unión a sus ligandos (MIP1 α , MIP1 β y RANTES), se comporta como un regulador clave en el tráfico de leucocitos y en la homeostasis del sistema inmune (Marzio PD *et al.*, 2005). *UBASH3B* codifica una proteína que contiene un dominio asociado a ubiquitina. Las reacciones de ubiquitinación se encuentran involucradas en la regulación de las señales de los receptores tirosina quinasa modulando la

endocitosis y la degradación de los receptores activados (Kowanzetz K *et al.*, 2004). Se ha demostrado que *UBASH3B* es un regulador negativo de la señal de los TCR (Carpino N *et al.*, 2004). *KIR2DL4* codifica una molécula KIR bastante peculiar desde el punto de vista funcional y mantiene una estrecha relación con las moléculas HLA de clase I.

Objetivo 3. Mejora del conocimiento del papel de las moléculas HLA en la enfermedad (trabajo 6).

Como estrategia para conseguir el **objetivo 3**, utilizamos los nuevos algoritmos de imputación desarrollados específicamente para la región HLA a partir de los datos obtenidos del iChip. Existen gran cantidad de datos que sugieren asociación de otras moléculas HLA con la enfermedad de manera independiente a la conocida y repetidamente comprobada asociación de *HLA-B*51*. Así, se ha descrito asociación con otros alelos *HLA-B* y numerosos estudios identifican alelos de riesgo en los otros genes *HLA* de clase I clásicos (*HLA-A* y *HLA-C*), algunos trabajos incluso atribuyen la asociación en la región HLA-B, no al propio gen *HLA-B* sino a variantes de un gen que codifica moléculas de clase I no clásicas, concretamente al gen *MICA*. (Montes-Cano MA *et al.*, 2013; Ombrello MJ *et al.*, 2014; Hughes T *et al.*, 2013). Existen una serie de condicionamientos que dificultan aclarar definitivamente este aspecto. Entre los más importantes se encuentra, el elevado grado de polimorfismo de cada uno de los genes *HLA* que dificulta la detección de todos los posibles alelos, que pueden ser desde muy frecuentes a muy raros y presentar diferencias poblacionales, lo que genera problemas estadísticos en la interpretación de resultados, ya que, los métodos de análisis masivos están más desarrollados para sistemas bialélicos. En relación con este hecho nos encontramos una serie de limitaciones, ya que, si existen múltiples alelos asociados a una enfermedad, una cohorte concreta tendrá mayor poder estadístico para aquellos alelos más comunes. Si, además, estos alelos presentan diferencias grandes de frecuencias entre poblaciones, en cada población se detectará asociación con el más común, por lo que, en principio, se

puede concluir que existen discrepancias entre diferentes trabajos, cuando realmente no las haya. Además, debemos tener en cuenta, que este elevado grado de polimorfismo afecta no sólo a un gen sino a muchos, ya que, existe gran densidad de genes que codifican moléculas relacionadas con el sistema inmune en esta región y que estos genes presentan un fuerte LD. Por último, en el caso concreto de BD, nunca se ha conseguido aclarar la razón biológica por la que determinadas moléculas HLA de clase I se relacionan con la enfermedad y el conocimiento de este aspecto está limitado también por la función dual que presentan muchas moléculas de clase I como presentadoras de péptidos a los linfocitos T CD8 y controladoras de la actividad NK por su función como ligando de los receptores KIR. Los nuevos sistemas de imputación (Jia X *et al.*, 2013; Browning BL *et al.*, 2009) nos permiten extraer mayor cantidad de información de la región HLA utilizando datos de genotipado masivo obtenidos en menor tiempo y de una manera más económica. Digamos que son métodos menos precisos para obtener el tipaje HLA de un individuo concreto, por lo que no podrían ser aplicados, por ejemplo, a los estudios que se realizan para trasplante, pero nos dan una información bastante cierta de las frecuencia alélicas en cohortes. Además, la información se puede manejar de diversas formas, a modo de alelos clásicos, SNPs y variantes de aminoácidos en las secuencias de proteínas.

La selección de la plataforma de genotipado es muy importante cuando se aborda un trabajo en el que se pretende imputar la región HLA, el iChip es la mejor plataforma para la imputación de la esta región, porque incluye muchos más SNPs que cualquier otro chip que se haya usado hasta la fecha para GWAS. Otro aspecto muy importante para realizar una buena imputación es el panel de referencia. Hay que tener en cuenta, que la correcta deducción de tipajes clásicos, aminoácidos, etc. basadas en el tipaje real de un conjunto determinado de SNPs depende de que el denominado panel de referencia, o sea, el conjunto de individuos tipados realmente para todos los marcadores, sea adecuado, o sea esté lo más próximo posible a la población de estudio desde el punto de vista genético. Existe un buen panel de referencia para población caucásica que incluye un gran

número de individuos, en nuestro trabajo utilizamos un panel que incluía datos de 5225 individuos de origen europeo referentes a 8961 SNPs e INDEL de la región HLA junto con los datos genotípicos de alelos HLA clase I y II con un nivel de resolución de cuatro dígitos (Brown WM *et al.*, 2009; de Bakker PI *et al.*, 2006). Trás la imputación, el conjunto final de datos a analizar incluía un total de 7227 SNPs, 419 alelos (125 alelos de 2 dígitos y 294 alelos de 4 dígitos) de los genes HLA de clase I (HLA-A, HLA-B, HLA-C) y clase II clásicos (HLA-DRB1, HLA-DQB1, HLA-DQA1, HLA-DPB1 and HLA-DPA1) y 1276 variantes aminoacídicas. Por eso, en el estudio de la región HLA no nos conformamos con la aplicación de los criterios de calidad generales para el iChip que se comentaron en relación con el objetivo 2, sino que comprobamos además la adecuación del panel y el análisis de imputación. Para ello, se compararon los datos obtenidos con dos métodos imputación, en el presente trabajo y métodos de tipaje tradicionales en un trabajo anterior de nuestro grupo (Montes-Cano *et al.*, 2013). Concretamente contábamos con el tipaje de grupos de alelos (2 dígitos) de los genes *HLA-A*, *HLA-B* y *HLA-C* de 278 individuos obtenidos por ambos método y cuantificamos la bondad de la imputación usando dos parámetros. Por un lado calculamos la precisión, “accuracy”, que es una medida de la veracidad de las frecuencias, el valor de este parámetro fue de alrededor al 98% en todos los casos, lo que nos indica que las frecuencias obtenidas por imputación, hablando en términos poblacionales, eran prácticamente idénticas a las obtenidas con métodos de tipaje tradicionales. Por otro lado, también queríamos comprobar la fiabilidad del tipaje de cada individuo y para ello utilizamos el valor del coeficiente de correlación (r). Como se puede observar en la **tabla S02 y en la figura S1 del trabajo 6**, los valores de r fueron, en general, muy buenos, superiores a 0.8 en prácticamente todos los casos, excepto para los grupos HLA-A*25 y HLA*26. En este caso, aunque las frecuencias poblacionales de ambos grupos de alelos coincidían en ambos métodos, los tipajes individuales no eran concordantes, por lo que decidimos eliminarlos de los análisis posteriores, ya que, en caso contrario podríamos llegar a conclusiones falsas en aquellos análisis que realizan estratificación por grupos de alelos como los multivariantes. Esto es un ejemplo de

porque no se puede pensar en la imputación como un sustituto de las técnicas de genotipado HLA tradicionales en el caso de situaciones que requieran una alta precisión de tipaje individual como sería en un trasplante, pero en cambio, es una técnica apropiada para estudios poblacionales de asociación a gran escala.

Los resultados del estudio de regresión logística de todos los marcadores de la región HLA, utilizando los 5 primeros componentes principales como covariable para controlar la posible estratificación de la población, indicaban que la señal con asociación más fuerte era la correspondiente al grupo de alelos HLA-B*51. Este resultado concuerda con todos los estudios previos hechos en BD y, de alguna manera, corrobora también la fiabilidad del análisis realizado. En cuanto al resto de loci, se descartó en toda la región de clase II, pero se detectó asociación con los otros 2 genes de clase I, *HLA-A* y *HLA-C*. En este sentido, nuestros resultados concuerdan con la mayoría de los publicados hasta la fecha que sugieren que podemos descartar que las moléculas de clase II estén involucradas en la predisposición a esta patología (Ombrello MJ *et al.*, 2014; Hughes T *et al.*, 2013). Para obviar, en la medida de lo posible, el problema del LD y determinar si las diferentes asociaciones encontradas eran independientes de HLA-B*51, realizamos un análisis de regresión logística condicional denominado “stepwise”. Nuestros datos, sustentan que la asociación del gen *HLA-A* es independiente de *HLA-B*, mientras que las asociaciones encontradas en el gen *HLA-C* son dependientes de *HLA-B*, es decir, se deben al LD entre ambos loci (**trabajo 6, tabla 1**). Estos resultados son concordantes con los descritos por Ombrello *et al.* y discrepan de los publicados por Hughes *et al.*, que sugieren que el gen *HLA-C* se encuentra asociado a BD de manera independiente de *HLA-B*. En cuanto a los resultados de este análisis, en el locus *HLA-B* encontramos asociados los grupos HLA-B*57 y HLA-B*15 que han sido previamente descritos en otras poblaciones. HLA-B*57 había sido descrito como factor de riesgo en un estudio realizado en población española (Montes-Cano MA *et al.*, 2014) y también en población británica (Ahmad T *et al.*, 2003) y HLA-B*15 se ha descrito en población marroquí (Choukri F *et al.*, 2001) y en población turca (Ombrello MJ *et al.*, 2014).

Con respecto a *HLA-A*, encontramos el grupo *HLA-A*02* como factor de riesgo independiente. En este sentido, se ha descrito asociación de *HLA-A*26* con la enfermedad en algunas poblaciones (Chung YM *et al.*, 1990; Kaburaki.T *et al.*, 2010). Este grupo de alelos fue eliminado de los análisis subsiguientes en nuestro estudio debido al bajo valor de *r* que presentaban los datos de la imputación. Sin embargo, los datos obtenidos mediante métodos de tipaje tradicional, no sugieren relación de este grupo alélico con la enfermedad en nuestra población. Además, tampoco se detectó asociación con *HLA-A*24*, que había sido descrito como factor de riesgo en un trabajo previo de nuestro grupo (Montes Cano *et al.*, 2013).

Para replicar la asociación de los genes *HLA-A* y *HLA-B*, genotipamos otra cohorte de pacientes y controles y, después de realizar los análisis correspondientes, se obtuvieron resultados similares (**trabajo 6, tabla S6 y tabla S7**). En este caso, que hay que resaltar que, el grupo de alelos *HLA-A*03* se encontró asociado a la enfermedad como factor de protección independiente. En la cohorte conjunta, resultado concordante con otros estudios (Ombrello MJ *et al.*).

Aunque es un hecho indiscutible que *HLA-B*51* se encuentra asociado con la enfermedad, sigue sin conocerse la causa de dicha asociación. Por ello la posibilidad de que la causa de la asociación resida en genes con elevado grado de LD como los genes *MICA* tuvo una gran acogida entre la comunidad científica. En 1999, se publicó un trabajo realizado en población japonesa que proponía que la asociación entre la enfermedad y la molécula *HLA-B51* se debía en realidad al LD de los genes *HLA-B* y *MICA* (Mizuki N *et al.*, 1999). Desde entonces se han llevado a cabo muchos estudios para reproducir esta asociación. Los resultados de la inmensa mayoría de estos estudios demuestran que la realidad era la contraria, es decir, que la asociación que estos autores describían entre la enfermedad y el gen *MICA* se debía a su fuerte LD con *HLA-B* (González-Escribano MF *et al.*, 1999; Yabuki K *et al.*, 1999; Wallace GR *et al.*, 1999; Mizuki N *et al.*, 2000, Salvarani C *et al.*, 2001; Cohen R *et al.*, 2002; Mizuki N *et al.*, 2002; Park SH *et al.*, 2002). No obstante, recientemente los resultados de dos nuevos estudios han vuelto a retomar esta idea de asociación de *MICA* con la enfermedad (Hughes T *et*

al., 2013; Carapito R *et al.*, 2015). Sin embargo, los resultados de nuestro trabajo y los de otro estudio a gran escala publicado recientemente coinciden con la gran mayoría de estudios de menor escala publicados anteriormente en que no existe una asociación independiente del gen *MICA* (Ombrello MJ *et al.*, 2014).

En nuestro caso, para intentar profundizar en el estudio de la región HLA realizamos un análisis que nos permitiera detectar que posiciones aminoacídicas están relacionadas con la enfermedad. Muchos aminoácidos de determinadas posiciones son comunes a proteínas codificadas por diferentes alelos HLA, por lo que, al centrarnos en el estudio de aminoácidos podemos obviar los problemas estadísticos derivados de la diversidad de frecuencias de los diferentes alelos y, además, conocer la región de la proteína responsable de la asociación. Con este propósito, utilizamos el test Omnibus que es un análisis de utilidad en el caso de variantes multi-alélicas, ya que, nos permite estudiar todas las diferentes variantes de estos polimorfismos. En el caso particular de los aminoácidos, nos permite estudiar la asociación de cada posición teniendo en cuenta todos los aminoácidos que se pueden encontrar en la posición en cuestión. Los resultados de este análisis en el que se incluyeron las diferentes posiciones aminoacídicas de las proteínas HLA de clase I y II, indican que la señal de asociación más fuerte corresponde con la posición 97 de la molécula HLA-B (**trabajo 6, tabla S7**) y que las posiciones de aminoácidos correspondientes a las moléculas de clase II, no muestran asociación. Después de condicionar por la posición 97 de HLA-B, para evitar los problemas derivados del LD, las señales de aminoácidos correspondientes a HLA-C desaparecen pero, sin embargo, se mantienen señales correspondientes HLA-A, siendo la de la posición 66 la que presenta la asociación más fuerte. Realizamos también un análisis de regresión logística con 5 PCA como covariables y un análisis “step-wise” de cada uno de los aminoácidos posibles en cada posición, obteniéndose resultados similares (**trabajo 6, tabla S10**). En la posición 97 de la molécula HLA-B podemos encontrar 6 aminoácidos, según nuestros resultados, Thr y Val serían factores de riesgo, Arg y Ser confieren protección y Asn y Trp se comportarían como neutros (**artículo 6, figura 2**). En cuanto a la posición 66 de la

molécula HLA-A, sólo hay posibilidad de 2 aminoácidos, Lys conferiría riesgo y Asn protección. Por tanto, el análisis de aminoácidos corrobora los resultados obtenidos con el análisis de grupos alélicos tradicionales, descartando asociación de la región de clase II y de la molécula HLA-C y sugiriendo asociaciones independientes de HLA-A y B y las posiciones más fuertemente relacionadas dentro de cada una de estas moléculas. Se hace necesaria una consideración especial con respecto a la interpretación de los análisis condicionados dentro de una misma molécula. Cada molécula HLA se caracteriza por tener un conjunto de aminoácidos determinados en las posiciones polimórficas, este conjunto de aminoácidos están, por definición, en LD y son dependientes. Por tanto, al condicionar por la posición más fuertemente asociada en una molécula concreta, la señal de las otras posiciones se pierde. Sin embargo, no creemos que, en este caso, podamos descartar con este tipo de análisis matemático la influencia de estas otras posiciones. Serían necesarios otro tipo de estudios, funcionales, de mutagénesis dirigida, etc. para determinar el conjunto de posiciones relevantes en una molécula concreta. Desde luego, este análisis señala las posiciones irrelevantes que serían aquellas que no presentan señal de asociación desde el principio o cuya señal se pierde al condicionar con posiciones relevantes de otra molécula.

Consideraciones finales. Las enfermedades de base inmunológica, entre las que se encuentra la enfermedad de Behçet, suponen un gran problema sanitario, ya que, tienen un impacto social y económico muy elevado. Se trata de enfermedades crónicas que afectan a la calidad de vida de los pacientes y generan problemas de incapacidad laboral e incluso, en los casos más graves, reducen la esperanza de vida de los pacientes. Como hemos comentado, muchas enfermedades inmuno-mediadas comparten manifestaciones clínicas, tratamientos, mecanismos moleculares y/o componente genético. El avance en el conocimiento de una de ellas, no solo supone ventajas para los pacientes de esa enfermedad en concreto, sino que nos ayuda a entender mejor el conjunto de todas. En este sentido, los resultados obtenidos en esta tesis doctoral confirman datos aportados por otros estudios y ofrecen nuevas pistas sobre las bases genéticas de la BD. Con el conocimiento actual, se puede empezar a esbozar un modelo de los mecanismos moleculares y rutas metabólicas implicados en la enfermedad (figura 12). Sustentando la hipótesis de que la enfermedad estaría desencadenada por determinados agentes ambientales y, entre los que los microorganismos podrían ejercer un papel primordial, los receptores encargados de detectar moléculas derivadas de estos microorganismos y las rutas que se desencadenan tras su activación representarían un papel importante en este hipotético modelo. Entre las rutas desencadenadas después de este reconocimiento, habría que destacar el inflammasoma, cuya activación supone un aumento de citoquinas proinflamatorias (IL1 β e IL18). Por otro lado, un papel de los microorganismos también puede estar relacionado con la ruta de procesamiento endógeno (sobre todo en caso de infecciones virales) y, en este sentido es donde cobrarían relevancia, el proteasoma (UBAHS3B) y las endopeptidasas del retículo endoplásmico (ERAP) que recortan los péptidos para que adquieran el tamaño adecuado para ser cargados a moléculas HLA de clase I. El heterodímero que constituye la molécula de clase I (cadena α polimórfica + β 2 microglobulina) no se exporta a superficie si no tiene un péptido cargado. Por tanto, la cantidad de péptidos generados por éstas enzimas, condiciona la expresión de moléculas HLA de clase I en superficie. Las moléculas HLA de clase I, como se ha explicado a lo largo de esta tesis, son los factores de

riesgo más relevantes en esta enfermedad. Su gran polimorfismo, determina que, cada molécula HLA sea capaz de unir con mayor afinidad un conjunto de péptidos diferentes. Por tanto, no solo la cantidad, sino también las características moleculares de los péptidos generados por ERAP pueden condicionar la expresión de moléculas HLA en superficie. A su vez, algunas moléculas HLA controlan la actividad NK mediante la interacción de los receptores KIR y presentan péptidos a los linfocitos T CD8, por lo que su papel en la enfermedad puede estar relacionado con cualquiera de estas funciones. Además, los linfocitos más relevantes en la enfermedad son los linfocitos T, y entre de ellos, los Th1 y Th17. La polarización de los linfocitos CD4 hacia Th1 o Th17 esta mediada por determinados pares interleucinas/receptor, algunas de las cuales se han relacionado con la enfermedad.

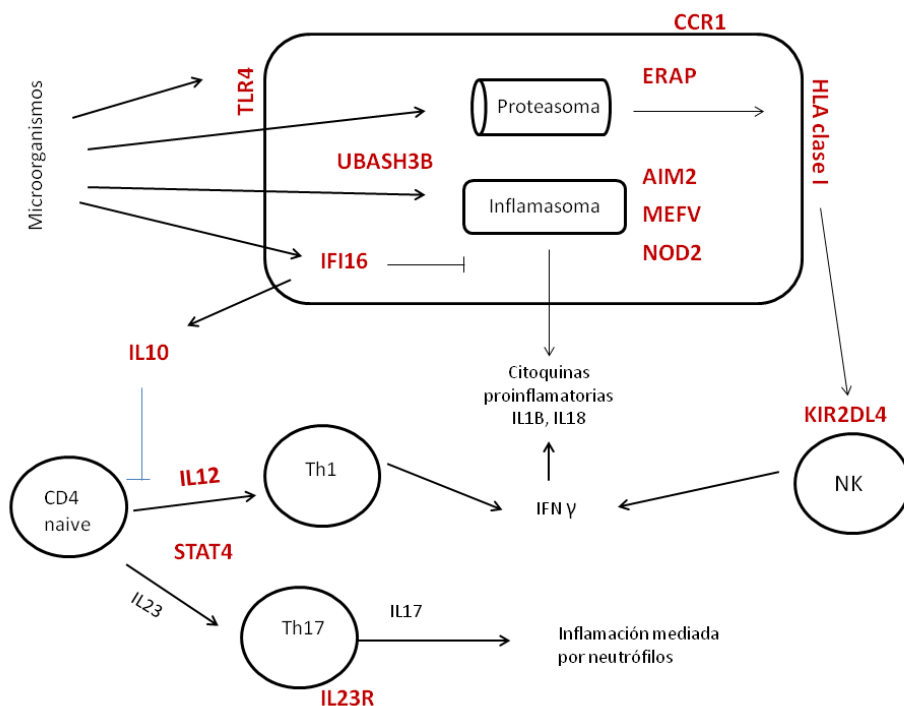


Figura 12. Modelo de rutas implicadas en la fisiopatología de la enfermedad de Behçet. Los genes que se encuentran asociados a la BD están en rojo

Perspectivas

En los últimos años se han alcanzado bastantes logros en el conocimiento de las bases genéticas de la enfermedad de Behçet, Así, nos encontramos en un punto en el que se puede comenzar a diseñar un modelo que integre las rutas moleculares implicadas en el desarrollo de la patología. Sin embargo, este modelo no está completo y aún queda camino por recorrer hasta llegar a dibujarlo completamente. Entre las causas que dificultan un avance más rápido en el conocimiento de este aspecto de la enfermedad, se encuentra su condición de enfermedad rara, que condiciona el tamaño de las cohortes de pacientes y, por tanto, el poder estadístico de los estudios individuales. Con esta limitación, se impone la necesidad de colaboración en estudios multicéntricos a nivel internacional que permitirían descifrar parte de la heredabilidad perdida y, también, investigar las diferencias poblacionales y analizar la influencia del componente genético en las diferentes características clínicas. En este sentido, las plataformas de genotipado de alto rendimiento, permiten análisis a posteriori de los datos obtenidos por cada grupo y requieren sólo del intercambio de datos, no de muestras biológicas ni de nuevos experimentos. Quizás, estemos en el momento de establecer este tipo de colaboración, ya que, hay varios estudios publicados con este tipo de plataformas en varias poblaciones. En este sentido, otra de las líneas que se beneficiarían de un estudio de este tipo, sería el de la asociación del sistema HLA con la enfermedad. Las nuevas aproximaciones para estudiar la región HLA que se centran en el estudio de los aminoácidos, pueden servir de base para plantear posibles estudios funcionales que ayuden a esclarecer el papel que desempeñan estas moléculas en la enfermedad. El siguiente paso sería llevar a cabo estudios funcionales que confirmen y expliquen la implicación de dichas rutas en la enfermedad.

Conclusiones

Conclusiones

1. Las asociaciones la región *GIMAP*, y los genes *TNFAIP3* y *JAK1* con la enfermedad de Behçet descritas en población asiática no se reproducen en población española, si bien, no se puede descartar completamente alguna influencia de estos genes en determinadas manifestaciones clínicas.
2. Con respecto a la variante *CCR5Δ32* y al gen *PTPN22*, los resultados en población española y los meta-análisis realizados descartan un papel relevante en la enfermedad de Behçet.
3. Nuestros resultados con respecto al rastreo de sensores de ácidos nucleicos sugieren asociación de los genes *AIM2/IFI16* con la enfermedad de Behçet. Esta asociación estaría relacionada con diferencias en los niveles de expresión de *IFI16* determinadas genéticamente que condicionarían diferentes ratios *AIM2/IFI16*.
4. Como resultado del rastreo de genes relacionados con el sistema inmune en la plataforma de alto rendimiento, además de replicar las asociaciones de los genes *IL23R*, *IL12A*, *STAT4*, *KIR2DL4*, *UBASH3B* y *CCR1* en población española, sugerimos una nueva región de asociación localizada entre los genes *JRKL* y *CNTN5*.
5. En cuanto a la región HLA, tanto en el análisis de grupos de alelos tradicionales como en el de aminoácidos se descarta asociación de las moléculas de clase II con la enfermedad. En cuanto a las moléculas de clase I, nuestros estudios sustentan asociación independiente de las moléculas HLA-A y B y descartan relación de las moléculas HLA-C y de otras moléculas HLA de clase I no clásicas como MICA.
6. Por último, el análisis de aminoácidos sugiere que las posiciones más relevantes en la enfermedad serían las posiciones 66 en la molécula HLA-A y la 97 de la molécula HLA-B.

Bibliografía

Adyunting Ao, Tokgoz G, D'Cruz DP, Gürler A, Cervera R, Düzgün N, *et al.* Antibodies to endothelial cells in patients with Behçet's disease. *Clin Immunol Immunopathol.* 1993; 67 (2): 157-62.

Ahmad T, Wallace GR, James T, Neville M, Bunce M, Mulcahy-Hawes K, *et al.* Mapping the HLA association in Behçet's disease: a role for tumor necrosis factor polymorphisms?. *Arthritis Rheum.* 2003; 48 (3): 807-13.

Atzeni F, Bioardi L, Casali B, Farnetti E, Nicoli D, Sarzi-Puttini P, *et al.* CC chemokine receptor 5 polymorphism in Italian patients with Behçet's disease. *Rheumatology.* 2012; 51 (12):2141-5.

Babbon JJ, Lucet IS, Murphy JM, Nicola NA, Varghese LN. The molecular regulation of Janus Kinase (JAK) activation. *Biochem J.* 2014; 462 (1):1-13.

Bank I, Duvdevani M, Livneh A. Expansion of gammadelta T-cells in Behçet's disease: role of disease activity and microbial flora in oral ulcers. *J Lab Clin Med.* 2003; 141 (1):33-40.

Baranathan V, Standford MR, Vaughan RW, Kondeatis E, Graham E, Fortune F, *et al.* The association of the PTPN22 620W polymorphism with Behçet's disease. *Ann Rheum Dis.* 2007; 66 (11): 1531-1533.

Behçet H. Some observations on the clinical picture of the so-called triple symptom complex. *Dermatologica.* 1940; 81: 73-78.

Behçet H. Über rezidivierende aphthose, durch ein virus verursachte Geshwuew am mund, am Auge, und an den Genitahen. *Dermatol Wochenschr.* 1937; 105:1152-7.

Behçet H. Considerations sue les lesions aphteuses de la bouche et des parties genitales, ainsi que sur les manifestations ocularies d'origine probablement virutique et observations concernant leur foyer d'infection. *Bull Soc Fr Dermatol Sympiligr.* 1938; 45: 420-433.

Ben Hamed M, Houman H, Milled M, Dellagi K, Louzir H. Involvement of chemokines and Th1 cytokines in pathogenesis of mucocutaneous lesions of Behçet's disease. *Arthritis Rheum.* 2004; 50 (7): 2291-5.

Bettencourt A, Leal B, Carvalho C Oliveira R, Martins Silva A, Vaz Patto J, *et al.* CC chemokine receptor polymorphism CCR5Δ32 in Portuguese Behçet's disease patients. *Clin Exp Rheumatol.* 2014; 32 (4 Suppl 84):S72-4.

Boone DL, Turer EE, Lee EG, Ahmad RC, Wheeler MT, Tsui C, *et al.* The ubiquitin-modifying enzyme A20 is required for termination of Toll-like receptor responses. *Nat Immunol.* 2004; 5 (10):1052-60.

Botstein D, Risch N. Discovering genotypes underlying human phenotypes: Past successes for mendelian disease, future approaches for complex disease. *Nat Genet.* 2003; 33:228-237.

Brown WM, Pierce J, Hilner JE, Perdue LH, Lohman K, Li L, *et al.* Overview of the MHC fine mapping data. *Diabetes Obes Metab.* 2009; 11 (Suppl 1):2-7.

Browning BL and Browning SR. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am J Hum Genet.* 2009; 84 (2): 889-899.

Burn GL, Svensson L, Sanchez-Blanco C, Saini M, Cope AP. Why is *PTPN22* a good candidate susceptibility gene for autoimmunity disease?. *FEBS Lett.* 2011; 585 (23): 3689-98.

Carapito R, Shahram F, Michel S, Le Gentil M, Radosavljevic M, Meguro A, *et al.* On the genetics of the Silk Route: association analysis of HLA, IL10, and IL23R-IL12RB2 regions with Behçet's disease in an Iranian population. *Immunogenetics.* 2015; 67 (5-6):289-293.

Cardon LR, Palmer LJ. Population stratification and spurious allelic association. *Lancet.* 2003; 361 (9357):598-604.

Carletto A, Pacor ML, Biasi D, Caramaschi P, Zeminian S, Bellavite P, *et al.* Changes of neutrophil migration without modification of in vitro metabolism and adhesion in Behçet's disease. *J Rheumatol.* 1997; 24 (7):1332-6.

Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet.* 2004; 74 (1):336-354.

Carmona FD, Mackie SL, Martín JE, Taylor JC, Vaglio A, Eyre S, *et al.* A large-scale genetic analysis reveals a strong contribution of the HLA class II region to giant cell arteritis susceptibility. *Am J Hum Genet.* 2015; 96 (4): 565-80.

Carpino N, Turner S, Mekala D, Takahashi Y, Zang H, Geiger TL, *et al.* Regulation of ZAP-70 activation and TCR signaling by two related proteins, Sts-1 and Sts-2. *Immunity.* 2004; 20 (1):37-46.

Cervera R, Navarro M, Lopez-Soto A, Cid MC, Font J, Esparza J, *et al.* Antibodies to endothelial cells in Behçet disease: cell-binding heterogeneity and association with clinical activity. *Ann Rheum Dis.* 1994; 53 (4): 265-7.

Chakravarti A. Population genetics—making sense out of sequence. *Nat Genet.* 1999; 21 (1 Suppl):56–60.

- Chang JT**, Shevach EM, Segal BM. Regulation of interleukin IL12 receptor beta2 subunit expression by endogenous IL12: a critical step in the differentiation of pathogenic autoreactive T cells. *J Exp Med*. 1999; 189 (6):969e978.
- Charo IF**, Ransohoff RM. The many roles of chemokines and chemokine receptors in inflammation. *N Engl J Med*. 2006; 354 (6):610-21.
- Chi W**, Zhou S, Yang P, Chen L. CD4+ T cells from behçet patients produce high levels of IL17. *Yan Ke Xue Bao*. 2011; 26 (2):6-13.
- Cho SB**, Cho S, Bang D. New insights in the clinical understanding of Behçet's disease. *Younsei Med J*. 2012; 53 (1):35-42.
- Choukri F**, Chakib A, Himmich H, H   S, Caillat-Zucman S.. HLA-B*51 and B*15 alleles confer predisposition to Behçet's disease in Moroccan patients. *Hum Immunol*. 2001; 62 (2):180-185.
- Chung YM**, Yeh TS, Sheu MM, Chen MS, Wen MS, Tsai HY, *et al*. Behçet's disease with ocular involvement in Taiwan: a joint survey of six major ophthalmological departments. *J Formos Med Assoc*. 1990; 89 (5): 413-7.
- Cohen R**, Metzger S, Nahir M, Chajek-Shaul T. Association of the MIC-A gene and HLA-B51 with Behçet's disease in Arabs and non-Ashkenazi Jewa in Israel. *Ann Rheum Dis*. 2002; 61 (2):157-60.
- Cortes A**, Brown MA. Promise and pitfalls of the Immunochip. *Arthritis Res Ther*. 2011; 13 (1):101.
- Dalghous AM**, Freysdottir J, Fortune F. Expression of cytokines, chemokines and chemokine receptors in oral ulcers of patients with Behçet's disease (BD) and recurrent aphtous stomatitis is Th1-associated, although Th2-association is also observed in patients with BD. *Scand J Rheumatol*. 2006; 35 (6): 472-5.
- de Bakker PI**, McVean G, Sabeti PC, Miretti MM, Green T, Marchini J, *et al*. A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat Genet*. 2006; 38 (10): 1166-72.
- de Bakker PI**, Raychaudhuri S. Interrogating the major histocompatibility complex with high-throughput genomics. *Hum Mol Genet*. 2012; 15 (R1): R29-36.
- de Menthon M**, Lavalley MP, Maldini C, Guillevin L, Mahr A.. HLA-B51/B5 and the risk of Behçet's disease: a systematic review and meta-analysis of case-control genetic association studies. *Arthritis Rheum*. 2009; 61 (10):1287-96.

Dieudé P, Guedj M, Wipff J, Avouac J, Hachulla E, Diot E, *et al.* The PTPN22 620W allele confers susceptibility to systemic sclerosis: findings of a large case-control study of European Caucasians and meta-analysis. *Arthritis Rheum Dis.* 2011; 58 (7): 2183-2188.

Dieudé P, Guedj M, Wipff J, Ruiz B, Riemekasten G, Matucci-Cerinic M, *et al.* Association of the TNFAIP3 rs5029939 variant with systemic sclerosis in the European Caucasian population. *Ann Rheum Dis.* 2010; 69 (11):1958-64.

Direskeneli H, Kenser G, D'Cruz DP, Khamashta MA, Akoğlu T, Yazici H, *et al.* Antiendothelial cells antibodies, endothelial proliferation von Willebrand factor antigen in Behçet's disease. *Clin Rheumatol.* 1995; 14 (1): 55-61.

Doria A, Zen M, Bettio S, Gatto M, Bassi N, Nalotto L, *et al.* Autoinflammation and autoimmunity: Briding the divide. *Autoimmun Rev.* 2012; 12 (1): 22-30.

Duan X, Ponomareva L, Veeranki S, Panchanathan R, Dickerson E, Choubey D. Differential roles for the interferon-inducible IFI16 and AIM2 innate immune sensors for cytosolic DNA in cellular senescence of human fibroblasts. *Mol Cancer Res.* 2011; 9 (5):589-602.

Düzgün N, Ayaşlıoğlu E, Tutkak H, Aydintuğ OT. Cytokine inhibitors: soluble tumor necrosis factor receptor 1 and interleukin-1 receptor antagonist in Behçet's disease. *Rheumatol Int.* 2005; 25 (1):1-5.

Efthimiou J, Addison IR, Johnson BV. In vivo leucocyte migration in Behçet's syndrome. *Ann Rheum Dis.* 1989; 48 (3):206-10.

Eksioglu-Demiralp E, Direskeneli H, Kibaroglu A, Yavuz S, Ergun T, Akoglu T. Neutrophil activation in Behçet's disease. *Clin Exp Rheumatol.* 2001; 19 (5 Suppl 24):S19-24.

Erlich H. HLA DNA typing: past, present, and future. *Tissue Antigens.* 2012; 80 (1):1-11.

Fei Y, Webb R, Cobb BI, Direskeneli H, Saruhan-Direskeneli G, Sawalha AH. Identification of novel genetic susceptibility loci for Behçet's disease using a genome-wide association study. *Arthritis Res Ther.* 2009; 11 (3):R66.

Ferrante A, Ciccia F, Principato A, Giardina AR, Impastato R, Peralta S, *et al.* A Th1 but not a Th17 response is present in the gastrointestinal involvement of Behçet's disease. *Clin Exp Rheumatol.* 2010; 28 (4 Suppl 60):S27-30.

Freigenbaum A. Description of Behçet's syndrome in the Hippocratic third book of endemic diseases. *Br J Ophtal.* 1956; 40 (6): 355-357.

Freysdottir J, Hussain L, Farmer I, Lau SH, Fortune F. Diversity of gammadelta T cells in patients with Behçet's disease is indicative of polyclonal activation. *Oral Dis.* 2006; 12 (3):271-7.

González-Escribano MF, Rodríguez MR, Aguilar F, Alvarez A, Sanchez-Roman J, Núñez-Roldán A. Lack of association of MICA transmembrane region polymorphism and Behçet's disease in Spain. *Tissue Antigens.* 1999; 54 (3):278-81.

González-Gay MA, García Porrúa C, Brañas F, López-Lázaro L, Olivieri I. Epidemiological and clinical aspects of Behçet's disease in a defined area of Northwestern Spain, 1988-1997. *J Rheumatol.* 2000; 27 (3):703-7.

Graña J, Sánchez-Meizoso MO, Galdo F. Epidemiological aspects of Behçet's disease in Galicia. *J Rheumatol.* 2001; 28 (11):2565-6.

Gül A, Tugal-Tutkum I, Dinarello CA, Reznikov L, Esen BA, Mirza A, *et al.* Interleukin-1 β -regulating antibody XOMA 052 (gevokizumab) in the treatment of acute exacerbations of resistant uveitis of Behçet's disease: an open-label pilot study. *Ann Theum Dis.* 2012; 71 (4): 563-6.

Gusella JF, Wexler NS, Conneal PM, Naylor SL, Anderson MA, Tanzi RE, *et al.* A polymorphic DNA marker genetically linked to Huntington's disease. *Nature.* 1983; 306 (5940):234-238.

Ha NT, Freytag S, Bikeboeller H. Coverage and efficiency in current SNP chips. *Eur J Hum Genet.* 2014; 22 (9):1124-30.

Hamzaoui K, Hamzaoui A, Hentati F, Kahan A, Ayed K, Chabbou A, *et al.* Phenotype and functional profile of T cells expressing gamma delta receptor from patients with Behçet's disease. *J Rheumatol.* 1994; 21 (12):2301-6.

Hamzaoui K, Kaabachi W, Fazaa B, Zakraoui L, Mili-Boussen I, Haj-Sassi F.. IL-33 levels and skin mRNA expression in Behçet's disease. *Clin Exp Rheumatol.* 2013; 31 (3 Suppl 77): 6-14.

Hattersley AT, McCarthy MI. What makes a good genetic association study?. *Lancet.* 2005; 366 (9493):1315-23.

Heath SC, Gut IG, Brennan P, McKay JD, Bencko V, Fabianova E, *et al.* Investigation of the fine structure of European populations with applications to disease association studies. *Eur J Hum Genet.* 2008; 16 (12):1413-1429.

Hong Li, Liu Q, Hou S, Du L, Zhou Q, Zhou Y, *et al.* *TNFAIP3* gene polymorphisms confer risk for Behçet's disease in a Chinese Han population. *Hum Genet.* 2013; 132 (3):293-300.

Hou S, Qi J, Zhang Q, Liao D, Li Q, Hu K, *et al.* Genetic variants in the *JAK1* gene confer higher risk of Behçet's disease with ocular involvement in Han Chinese. *Hum Genet.* 2013; 132 (9):1049-1058.

Hou S, Yang Z, Du L, Jiang Z, Shu Q, Chen Y, *et al.* Identification of a susceptibility locus in *STAT4* for Behçet's disease in Han Chinese in a Genome-wide association study. *Arthritis Rheum.* 2012; 64 (12):4104-13.

Hu K, Hou S, Li F, Xiang Q, Kijlstra A, Yang P. *JAK1*, but not *JAK2* and *STAT3*, confers susceptibility to Vogt-Koyanagi-Hatada (VKH) syndrome in a Han Chinese population. *Invest Ophthalmol Vis Sci.* 2013; 54 (5):3360-5.

Hughes T, Coir P, Alder A, Yilmaz V, Aksu K, Düzgün N, *et al.* Identification of multiple independent susceptibility loci in the HLA region in Behçet's disease. *Nat Genet.* 2013; 45 (3):319-24.

International Study group for Behcet's disease. Criteria for diagnosis of Behcet's disease. *Lancet.* 1990; 335:1078-1080.

Ionnidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. *Nat Genet.* 2001; 29 (3):306-309.

Iwakura Y, Ishigame H. The IL23/IL17 axis in inflammation. *J Clin Invest.* 2006; 116 (5): 1218-22.

Jia X, Han B, Onengut-Gumuscu S, Chen WM, Concannon PJ, Rich SS, *et al.* Imputing amino acid polymorphisms in human leukocyte antigens. *PLoS ONE.* 2013; 8 (6):e64683.

Jiang Z, Yang P, Hou S, Du L, Xie L, Zhou H, *et al.* IL-23R gene confers susceptibility to Behcet's disease in a Chinese Han population. *Ann Rheum Dis.* 2010; 69 (7):1325-8.

Kanenko F, Oyama N, Yanagihori H, Isogai E, Yokota K, Oguma K. The role of streptococcal hypersensitivity in the pathogenesis of Behçet's disease. *Eur J Dermatol.* 2008; 18 (5):489-498.

Kang EH, Choi JY, Lee YJ, Lee EY, Lee EB, Song YW. Single nucleotide polymorphisms in IL-10-mediated signalling pathways in Korean patients with Behçet's disease. *Clin Exp Rheumatol.* 2014; 32 (4 Suppl 84):S27-32.

Kappen JH, Medina-Gomez C, van Hagen PM, Stolk L, Estrada K, Rivadeneira F, *et al.* Genome-wide association studies in an admixed case series reveals IL12A as a new candidate in Behçet's disease. *PLoS ONE.* 2015; 10 (3):e0119085.

Kapsimali VD, Kanakis MA, Vaiopoulos GA, Kaklamanis PG. Etiopathogenesis of Behçet's disease with emphasis on the role of immunological aberrations. *Clin Rheumatol.* 2010; 29 (11):1211-1216.

- Karubaki T**, Genetic association of HLA-A*2601 with ocular Behçet's disease in Japanese patients. *Clin Exp Rheumatol*. 2010; 28 (4 Suppl 60): S39-44.
- Kawakami T**, Yamazaki M, Mizoguchi M, Soma Y. Antiphosphatidylserine-prothrombin complex antibodies in 3 patients with Behçet disease involving superficial vein thrombophlebitis. *Arch Dermatol*. 2009; 145 (2):171-5.
- Kaya TI**. Genetics of Behçet's disease. *Patholog Res Int*. 2012; 2012:912589.
- Kim ES**, Kim SW, Moon CM, Park JJ, Kim TI, Kim WH, *et al*. Interactions between IL17A, IL23R, and STAT4 polymorphisms confer susceptibility to intestinal Behçet's disease in Korean population. *Life Sci*. 2012; 90 (19-20):740-6.
- Kim J**, Park JA, Lee EY, Lee YJ, Song YW, Lee EB. Imbalance of Th17 to Th1 cells in Behçet's disease. *Clin Exp Rheumatol*. 2010; 28 (4 Suppl 60):S16-9.
- Kirino Y**, Bertsias G, Ishigatsubo Y, Mizuki N, Tugal-Tutkun I, Seyahi E, *et al*. Genome-wide association analysis identifies new susceptibility loci for Behçet's disease and epistasis between HLA-B*51 and ERAP1. *Nat Genet*. 2013; 45 (2):202-207.
- Kirino Y**, Zhou Q, Ishigatsubo Y, Mizuki N, Tugal-Tutkun I, Seyahi E, *et al*. Targeted resequencing implicates the familial Mediterranean fever gene MEFV and the toll-like receptor 4 gene TLR4 in Behçet disease. *Proc Natl Acad Sci U S A*. 2013; 110 (20):8134-9.
- Kobayashi M**, Ito M, Nakagawa A, Matsushita M, Nishikimi N, Sakurai T, *et al*. Neutrophil and endothelial cell activation in the vasa vasorum in vasculo-Behçet disease. *Histopathology*. 2004; 36 (4):362-71.
- Koné-Paut I**, Geisler I, Wechsler B, Ozen S, Ozdogan H, Rozenbaum M, *et al*. Familial aggregation in Behçet's disease: high frequency in siblings and parents of pediatric probands. *J Pediatr*. 1999; 135 (1):89-93.
- Kowanetz K**, Crosetto N, Haglund K, Schmidt MH, Heldin CH, Dikic I. Suppressors of T-cell receptor signaling Sts-1 and Sts-2 bind to Cbl and inhibit endocytosis of receptor tyrosine kinases. *J Biol Chem*. 2004, 279 (31):32786-32795.
- Lander ES**, Schork NJ. Genetic dissection of complex traits. *Science*. 1994; 265 (5181):2037-2048.
- Lander ES**. The new genomics: global views of biology. *Science*. 1996; 274 (5287):536-539.
- Lao O**, Lu TT, Nothnagel M, Junge O, Freitag-Wolf S, Caliebe A, *et al*. Correlation between genetic and geographic structure in Europe. *Curr Biol*. 2008; 18 (16):1241-1248.

Lee S, Bang D, Cho YH, Lee ES, Sohn S. Polymerase chain reaction reveals herpes simplex virus DNA in saliva of patients with Behçet's disease. *Arch Dermatol Res.* 1996; 288 (4):179-83.

Lee YJ, Horie Y, Wallace GR, Choi YS, Park JA, Choi JY, *et al.* Genome-wide association study identifies GIMAP as a novel susceptibility locus for Behçet's disease. *Ann Rheum Dis.* 2013; 72 (9): 1510-1516.

Lewallen S, Courtright P. Epidemiology in practice: case-control studies. *Community Eye Health.* 1998; 11 (28):57-58.

Lewontin RC, Kojima K. The evolutionary dynamics of complex polymorphisms. *Evolution* 1960; 14:458-472.

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, *et al.* Finding the missing heritability of complex diseases. *Nature.* 2009; 461 (7265):747-753.

Marchini J, Howie B. Genotype imputation for genome-wide association studies. *Nat Rev Genet.* 2010; 11 (7):499-511.

Martinez A, Valdivia A, Pascual-Salcedo D, LamasJR, Fernandez-Arquero M, Balsa A, *et al.* PADI4 polymorphisms are not associated with rheumatoid arthritis in the Spanish population. *Rheumatology* 2005;44:1263-6.

Marzio PD, Dai WW, Franchin G, Chan AY, Symons M, Sherry B. Role of Rho family GTPases in CCR1 and CCR5-induced actin reorganization in macrophages. *Biochem Biophys Res Commun.* 2005; 331 (4):909-916.

Masatlioglu S, Seyahi E, Tahir Turanli E, Fresko I, Gogus F, Senates E, *et al.* A twin study in Behçet's syndrome. *Clin Exp Rheumatol.* 2010; 28 (4 Suppl 60):S62-6.

Matos M, Xavier JM, Abrantes P, Sousa I, Rei N, Davatchi F, *et al.* IL10 low-frequency variants in Behçet's disease patients. *Int J Rheum Dis.* 2014 Apr 8. doi: 10.1111/1756-185X.

Mayes MD, Bossini-Castillo L, Gorlova O, Martin JE, Zhou X, Chen WV, *et al.* ImmunoChip analysis identifies multiple susceptibility loci for systemic sclerosis. *Am J Hum Genet.* 2014; 94 (1):47-61.

McGonagle D, McDermott MF. A proposed classification of the immunological diseases. *PLoS Med.* 2006; 3 (8): e297.

Mendes D, Correia M, Barbedo MJ, Vaio T, Mota M, Gonçalves O, *et al.* Behçet's disease: a contemporary review. *J Autoimmun.* 2009; 32 (3-4):178-188.

- Miretti MM**, Walsh EC, Ke X, Delgado M, Griffiths M, Hunt S, *et al.* A high-resolution linkage-disequilibrium map of the human major histocompatibility complex and first generation of tag single-nucleotide polymorphisms. *Am J Hum Genet.* 2005; 76 (4):634-46.
- Mizuki N**, Meguro A, Ota M, Ohno S, Shiota T, Kawagoe T, *et al.* Genome-wide association study identify IL23R-IL12RB2 and IL10 as Behçet's disease susceptibility loci. *Nat Genet.* 2010; 42 (8):703-6.
- Mizuki N**, Ota M, Katsuyama Y, Yabuki K, Ando H, Goto K, *et al.* Association analysis between the MIC-A and HLA-B alleles in Japanese patients with Behçet's disease. *Arthritis Rheum.* 1999; 42 (9): 1961-6.
- Mizuki N**, Ota M, Katsuyama Y, Yabuki K, Ando H, Shiina T, *et al.* Sequencing-based typing of HLA-B*51 alleles and the significant association of HLA-B5101 and HLA-B5108 with Behçet's disease in Greek patients. *Tissue Antigens.* 2002; 59 (2):118-21.
- Mizuki N**, Ota M, Yabuki K, Katsuyama Y, Ando H, Palimeris GD, *et al.* Localization of the pathogenic gene of EBhçet's disease by microsatellite analysis of three different populations. *Invest Ophthalmol Vis Sci.* 2000; 41 (12):3702-8.
- Mizushima Y**, Matsuda T, Hoshi K, Ohno S. Induction of Behçet's disease symptoms after dental treatment and streptococcal antigen skin test. *J Rheumatol.* 1988; 15 (6): 1029-30.
- Mojtahedi Z**, Ahmadi SB, Razmkhah M, Azad TK, Rajaei A, Ghaderi A. Association of chemokine receptor 5 (CCR5) Δ 32 mutation with Behçet's disease is dependent on gender in Iranian patients. *Clin Exp Rheumatol.* 2006; 24 (5 Suppl 42):S91-94.
- Monselise A**, Weinberger A, Monselise Y, Fraser A, Sulkes J, Krause I. Anti-Saccharomyces cerevisiae antibodies in Behçet's disease- a family study. *Clin Exp Rheumatol.* 2006; 24 (5 Suppl 42):S87-90.
- Montes-Cano MA**, Conde-Jaldón M, García-Lozano JR, Ortiz-Fernández L, Ortego-Centeno N, Castillo-Palma MJ, *et al.* HLA and non-HLA genes in Behçet's disease: a multicentric study in the Spanish population. *Arthritis Res Ther.* 2013; 15 (5): R145.
- Morinobu A**, Gadina M, Strober W, Visconti R, Fornace A, Montagna C, *et al.* STAT4 serine phosphorylation is critical for IL-12 induced IFN-gamma production but not for cell proliferation. *Proc Natl Acad Sci USA* 2002; 99 (19):12281-12286.
- Musabak U**, Baylan O, Cetin T, Yesilova Z, Sengul A, Saglam K, *et al.* Lipid profile and anticardiolipin antibodies in Behçet's disease. *Arch Med Res.* 2005; 36 (4): 387-92.

Musabak U, Pay S, Erdem H, Simsek I, Pekel A, Dinc A, *et al.* Serum interleukin-18 levels in patients with Behçet's disease. Is its expression associated with disease activity or clinical presentations?. *Rheumatol Int.* 2006; 26 (6):545-50.

Neves FS, Carrasco S, Goldenstein-Schainberg, Gonçalves CR, de Mello SB. Neutrophil hyperchemotaxis in Behçet's disease: a possible role for monocytes orchestrating bacterial-induced innate immune responses. *Clin Rheumatol.* 2009; 28 (12):1403-10.

Ombrello MJ, Kirino Y, de Bakker PI, Gül A, Kastner DL, Remmers EF. Behçet disease-associated MHC class I residues implicate antigen binding and regulation of cell-mediated cytotoxicity. *Proc Natl Acad Sci USA.* 2014; 111 (24):8867-72.

Ono S, Aoki K, Sugiura S, Nakayama E, Itakura K. Letter: HL-A5 and Behçet's disease. *Lancet.* 1973; 2 (7842):1383-4.

Park SH, Park Ks, Seo YI, Min DJ, Kim WU, Kim TG, *et al.* Association of MICA polymorphism with HLA-B51 and disease severity in Korean patients with Behçet's disease. *J Korean Med Sci.* 2002; 17 (3):336-70.

Pitchard JK. Are rare variants responsible for susceptibility to complex diseases? *Am J Hum Genet.* 2001; 69 (1):124-137.

Reich D, Thangaraj K, Patterson N, Price AL, Singh L. Reconstructing Indian population history. *Nature.* 2009; 461 (7263):489-494.

Remmers EF, Cosan F, Kirino Y, Ombrello MJ, Abaci N, Satorius C, *et al.* Genome-wide association study identifies variants in the MHC class I, IL10, and IL23R-IL12RB2 regions associated with Behçet's disease. *Nat Genet.* 2010; 42 (8):698-702.

Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science.* 1996; 273 (5281):1516-1517.

Robinson J, Halliwell JA, Hayhurst JD, Flicek P, Parham P, Marsh SG. The IPD and IMGT/HLA database: allele variant database. *Nucleic Acid Res.* 2015; 43:D423-31.

Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky LA, *et al.* Genetic structure of human populations. *Science.* 2002; 298 (5602):2381-2385.

Sahin N, Bicakcigil M, Atagunduz P, Direskeneli H, Saruhan-Direskeneli G. *PTPN22* gene polymorphism in Behçet's disease. *Tissue Antigens.* 2007; 70 (5): 432-4.

- Salvarani C**, Boiardi L, Mantocani V, Olivieri I, Ciancio G, Cantini F, *et al.* Association of MICA alleles and HLA-B51 in Italian patients with Behçet's disease. *J Rheumatol.* 2001; 28 (8):1867-70.
- Saunders A**, Webb LM, Janas ML, Hutchings A, Pascall J, Carter C, *et al.* Putative GTPase GIMAP1 is critical for the development of mature B and T lymphocytes. *Blood.* 2010; 115 (16):3249-57.
- Sham PC**, Purcell SM. Statistical power and significance testing in large-scale genetic studies. *Nat Rev Genet.* 2014; 15 (5):335-346.
- Shon S**. Etiopathology of Behçet's disease; Herpes simplex virus infection and animal model. *Yonsei Med J.* 1997; 38 (6):359-363.
- Sohn S**, Lee ES, Bang D, Lee S. Behçet's disease-like symptoms induced by the Herpes simplex virus in ICR mice. *Eur J Dermatol.* 1998; 8 (1): 21-3.
- Stanford SM**, Bottini N. *PTPN22*: the archetypal non-HLA autoimmunity gene. *Nat Rev Rheumatol.* 2014; 10 (10): 602-11.
- Suzuki A**, Yamada R, Chang X, Tokuhira S, Sawada T, Suzuki M, *et al.* Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. *Nat Genet* 2003;34:395-402.
- Tabor HK**, Risch N, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat Rev Genet.* 2002; 3 (5): 391-7.
- Tasliyurt T**, Yigit S, Rustemoglu A, Gul U, Ates O. Common MEFV gene mutations in Turkish patients with Behçet's disease. *Gene.* 2013; 530 (1):100-3.
- Tenesa A**, Dunlop MG. Validating of tagging SNPs across populations for association studies. *Eur J Hum Genet.* 2006; 14 (3):337-345.
- The 1000 Genomes Project Consortium**. An integrated map of genetic variation from 1,092 human genomes. *Nature.* 2012; 491 (7422): 56-65.
- The HUGO Pan-Asian SNP Consortium**. Mapping human genetic diversity in Asia. *Science* 2009; 326 (5959):1541-1545.
- The International HapMap Consortium**. A haplotype map of the human genome. *Nature* 2005; 437:1299-1320
- Tiwari HK**, Barnholtz-Sloan J, Wineinger N, Padilla MA, Vaughan LK, Allison DB. Review and evaluation of methods for correcting for population stratification with a focus on underlying statistical principles. *Hum Hered.* 2008; 66 (2): 67-86.

Traherne JA. Human MHC architecture and evolution: implications for disease association studies. *Int J Immunogenet.* 2008; 35 (3):179-92.

Veeranki S, Duan X, Panchanathan R, Liu H, Choubey D. IFI16 protein mediates the anti-inflammatory actions of the type-I interferons through suppression of activation of caspase-1 by inflammasomes. *PLoS One.* 2011; 6 (10):e27040.

Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS Discovery. *Am J Hum Genet.* 2012; 90 (1):7-24.

Wallace GR, Verity DH, Delamaine LJ, Ohno S, Inoko H, Ota M, *et al.* MIC-A allele profiles and HLA class I associations in Behçet's disease. *Immunogenetics.* 1999; 49 (7-8):613-7.

Wu Z, Zheng W, Xu J, Sun F, Chen H, Li P, *et al.* IL10 polymorphisms associated with Behçet's disease in Chinese Han. *Hum Immunol.* 2014 ; 75 (3):271-6.

Xavier JM, Shahram F, Davatchi F, Rosa A, Crespo J, Abdollahi BS, *et al.* Association study of IL10 and IL23R-IL12RB2 in Iranian patients with Behçet's disease. *Arthritis Rheum.* 2012; 64 (8):2761-72.

Yabuki K, Mizuki N, Ota M, Katsuyama Y, Palimeris G, Stavropoulos C, *et al.* Association of MICA gene and HLA-B*5101 with Behçet's disease in Greece. *Invest Ophthalmol Bis Sci.* 1999; 40 (9):1921-6.

Yalçın B, Atakan N, Dogan S. Association of interleukin-23 receptor gene polymorphism with Behçet disease. *Clin Exp Dermatol.* 2014; 39 (8):881-7.

Yang X, Ahmad T, Gogus F, Verity D, Wallace GR, Madanat W, *et al.* Analysis of the CC chemokine receptor 5 (CCR5) Δ32 polymorphism in Behçet's disease. *Eur J Immunogenet.* 2004; 31 (1):11-4.

Yazici H, Fresko I, Yurdakul S. Behçet's syndrome: disease manifestations, management, and advances in treatment. *Nat Clin Pract Rheumatol.* 2007; 3 (3):148-155.

Yurdakul S, Yazici H. Behçet syndrome. *Best Pract Res Clin Rheumatol* 2008; 22: 793-809.

Zhang M, Xu WD, Wen PF, Liang Y, Liu J, Pan HF, *et al.* Polymorphisms in the tumor necrosis factor gene and susceptibility to Behcet's disease: an updated meta-analysis. *Mol Vis.* 2013; 19:1913-24.

Zhang Q, Hou S, Jiang Z, Du L, Li F, Xiao X, *et al.* No association of PTPN22 polymorphisms with susceptibility to ocular Behçet's disease in two Chinese Han populations. *PLoS ONE* 2012; 7 (3): e31230.

Zhang X, Li W, Zhang X, Zhao L, Zhang X, Jiang L, *et al.* Single nucleotide polymorphisms in TNFAIP3 were associated with the risks of rheumatoid arthritis in northern Chinese Han population. BMC Medical Genetics. 2014; 15: 15-56.

Ziv E, Burchard Eg. Human population structure and genetic association studies. Pharmacogenomics. 2003; 4 (4): 431-441.

Anexos

Otras publicaciones del doctorando

Ortiz-Fernández L, J. García-Lozano, M. Montes-Cano, M. Conde-Jaldón, J. Márquez, E. Leo, M. Gómez-García, N. Ortego-Centeno, F. García-Hernández, G. Espinosa, G. Graña-Gil, J. Sánchez-Bursón, M. Juliá, R. Blanco, A. Barnosi-Marín, R. Solans, P. Fanlo, M. Rodríguez-Carballeira, T. Camps, S. Castañeda, J. Martín, M. González-Escribano, **Association of haplotypes of the TLR8 locus with susceptibility to Crohn's and Behçet's disease**. Clin Exp Rheum 2015. Aceptado para publicación.

Carmona FD, Mackie SL, Martín JE, Taylor JC, Vaglio A, Eyre S, Bossini-Castillo L, Castañeda S, Cid MC, Hernández-Rodríguez J, Prieto-González S, Solans R, Ramentol-Sintas M, González-Escribano MF, **Ortiz-Fernández L**, Morado IC, Narváez J, Miranda-Filloy JA; Spanish GCA Group, Beretta L, Lunardi C, Cimmino MA, Gianfreda D, Santilli D, Ramirez GA, Soriano A, Muratore F, Pazzola G, Addimanda O, Wijmenga C, Witte T, Schirmer JH, Moosig F, Schönau V, Franke A, Palm Ø, Molberg Ø, Diamantopoulos AP, Carette S, Cuthbertson D, Forbess LJ, Hoffman GS, Khalidi NA, Koenig CL, Langford CA, McAlear CA, Moreland L, Monach PA, Pagnoux C, Seo P, Spiera R, Sreih AG, Warrington KJ, Ytterberg SR, Gregersen PK, Pease CT, Gough A, Green M, Hordon L, Jarrett S, Watts R, Levy S, Patel Y, Kamath S, Dasgupta B, Worthington J, Koeleman BP, de Bakker PI, Barrett JH, Salvarani C, Merkel PA, González-Gay MA, Morgan AW, Martín J. **A large-scale genetic analysis reveals a strong contribution of the HLA class II region to giant cell arteritis susceptibility**. Am J Hum Genet. 2015; 96:565-80.

López-Mejías R, Genre F, Pérez BS, Castañeda S, Ortego-Centeno N, Llorca J, Ubilla B, Remuzgo-Martínez S, Mijares V, Pina T, Calvo-Río V, Márquez A, Miranda-Filloy JA, Parejo AN, Conde-Jaldón M, **Ortiz-Fernández L**, Argila D, Aragües M, Rubio E, Luque ML, Blanco-Madriral JM, Galíndez-Aguirregoikoa E, González-Escribano F, Ocejó-Vinyals JG, Martín J, Blanco R, González-Gay MA. **Association of HLA-B*41:02 with Henoch-Schönlein Purpura (IgA Vasculitis) in Spanish individuals irrespective of the HLA-DRB1 status**. Arthritis Res Ther. 2015; 14:17:102.

López-Mejías R, Genre F, Pérez BS, Castañeda S, Ortego-Centeno N, Llorca J, Ubilla B, Remuzgo-Martínez S, Mijares V, Pina T, Calvo-Río V, Márquez A, Salazar-Icardo L, Miranda-Filloy JA, Conde-Jaldón M, **Ortiz-Fernández L**, Rubio E, León Luque M, Blanco-Madriral JM, Galíndez-Aguirregoikoa E, González-Vela MC, Ocejó-Vinyals JG, González-Escribano F, Martín J, Blanco R, González-Gay MA. **HLA-DRB1 association with Henoch-Schönlein purpura**. Arthritis Rheumatol. 2014. doi: 10.1002/art.38979.

Conde-Jaldón M, Montes-Cano MA, García-Lozano JR, *Ortiz-Fernández L*, Ortego-Centeno N, González-León R, Espinosa G, Graña-Gil G, Sánchez-Bursón J, González-Gay MA, Barnosi-Marín AC, Solans R, Fanlo P, Carballeira MR, Camps T, Castañeda S, Martín J, González-Escribano MF. **Epistatic interaction of ERAP1 and HLA-B in Behçet disease: a replication study in the Spanish population.** PLoS One. 2014. 14; 9:e102100.

Montes-Cano MA, Conde-Jaldón M, García-Lozano JR, *Ortiz-Fernández L*, Ortego-Centeno N, Castillo-Palma MJ, Espinosa G, Graña-Gil G, González-Gay MA, Barnosi-Marín AC, Solans R, Fanlo P, Camps T, Castañeda S, Sánchez-Bursón J, Núñez-Roldán A, Martín J, González-Escribano MF. **HLA and non-HLA genes in Behçet's disease: a multicentric study in the Spanish population.** Arthritis Res Ther. 2013. 4; 15:R145.

García-Lozano JR, Torres-Agrela B, Montes-Cano MA, *Ortiz-Fernández L*, Conde-Jaldón M, Teruel M, García A, Núñez-Roldán A, Martín J, González-Escribano MF. **Association of the AIRE gene with susceptibility to rheumatoid arthritis in a European population: a case control study.** Arthritis Res Ther. 2013. 15; 15:R11.

Comunicaciones orales a congresos como primer firmante

Aplicación de la tecnología Sequenom: Escrutinio de genes candidatos en rutas de inmunidad innata. Ortiz-Fernández L. X Foro de Investigadores HUVR/IBiS, “La investigación clínica en el campus HUVR-IBiS” Sevilla, 14-15 diciembre 2012

Asociación de las proteínas Ifi-200 (IFI16 y AIM2) con la enfermedad de Behçet. L Ortiz-Fernández, JR García-Lozano, MA Montes-Cano, M Conde-Jaldón, N Ortego-Centeno, MJ Castillo-Palma, G Espinosa, G Graña-Gil, MA González-Gay, AC Barnosi-Marín, R Solans, P Fanlo, T Camps, S Castañeda, J Sánchez-Bursón, A Núñez-Roldán, J Martín, MF González-Escribano. XXXVII Congreso de la Sociedad Española de Inmunología. Salamanca 29-31 Mayo/ 1 de Junio 2013

Asociación de las proteínas Ifi-200 (IFI16 y AIM2) con la enfermedad de Behçet. L Ortiz-Fernández, JR García-Lozano, MA Montes-Cano, M Conde-Jaldón, N Ortego-Centeno, MJ Castillo-Palma, G Espinosa, G Graña-Gil, MA González-Gay, AC Barnosi-Marín, R Solans, P Fanlo, T Camps, S Castañeda, J Sánchez-Bursón, A Núñez-Roldán, J Martín, MF González-Escribano. XXXIV Congreso Nacional de la Sociedad Española de Medicina Interna. Málaga 21-23 Noviembre 2013.

Identificación del cluster AIM2-IFI16 como un nuevo locus de susceptibilidad en la enfermedad de Behçet utilizando un screening de sensores de ácidos nucleicos. L Ortiz-Fernández, JR García-Lozano, MA Montes-Cano, M Conde-Jaldón, N Ortego-Centeno, MJ Castillo-Palma, G Espinosa, G Graña-Gil, MA González-Gay, AC Barnosi-Marín, R Solans, P Fanlo, T Camps, S Castañeda, J Sánchez-Bursón, A Núñez-Roldán, J Martín, MF González-Escribano. VII Congreso de la Asociación Andaluza de Enfermedades Autoinmunes. Almería 13-15 Febrero 2015.